

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 9 08:36:34 1997; MasPar time 215.75 Seconds
Tabular output not generated. 980.525 Million cell updates/sec

Title: >US-08-848-439-1
Description: (1-2027) from US08848439.seq
Perfect Score: 2027
N.A. Sequence: 1 GAATTCGGCCTTCATGGCCT.....AAAAAAGCGGCCGC 2027
Comp: CTTAAGCCGGAAGTACCGA.....TTTTTTTTCGCCGCGCG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0
Searched: 142080 seqs, 52183452 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq28
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29

Statistics: Mean 9.918; Variance 6.937; scale 1.430

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	201	9.9	204	20	T23535 Human gene signature	1.17e-97
2	88	4.3	1047	2	Q10572 Human Natriuretic Pep	3.43e-32
3	84	4.1	1047	2	Q10572 Human Natriuretic Pep	5.69e-30
4	44	2.2	91	9	Q51746 Oligonucleotide probe	8.73e-09
5	43	2.1	204	1	N81164 Base substituted E.co	2.71e-08
6	41	2.0	91	9	Q51746 Oligonucleotide probe	2.56e-07
7	40	2.0	114	12	Q70468 Generic DNA sequence	7.77e-07
8	40	2.0	114	12	Q70466 Generic DNA sequence	7.77e-07
9	40	2.0	204	1	N81164 Base substituted E.co	7.77e-07
10	39	1.9	114	12	Q70469 Generic DNA sequence	2.34e-06
11	38	1.9	114	12	Q70465 Generic DNA sequence	7.00e-06
12	38	1.9	114	12	Q70467 Generic DNA sequence	7.00e-06
13	39	1.9	114	12	Q70469 Generic DNA sequence	2.34e-06
14	38	1.9	114	12	Q70466 Generic DNA sequence	7.00e-06
15	38	1.9	114	12	Q70468 Generic DNA sequence	7.00e-06

16	36	1.8	114	12	Q70470	Generic DNA sequence	6.09e-05
17	36	1.8	114	12	Q70472	Generic DNA sequence	6.09e-05
18	36	1.8	114	12	Q70467	Generic DNA sequence	6.09e-05
19	34	1.7	114	12	Q70473	Generic DNA sequence	5.10e-04
20	34	1.7	114	12	Q70471	Generic DNA sequence	5.10e-04
21	34	1.7	114	12	Q70465	Generic DNA sequence	5.10e-04
22	34	1.7	114	12	Q70471	Generic DNA sequence	5.10e-04
23	34	1.7	114	12	Q70470	Generic DNA sequence	5.10e-04
24	32	1.6	114	12	Q70472	Generic DNA sequence	4.09e-03
25	30	1.5	565	6	Q35072	HCV envelope region n	3.13e-02
26	30	1.5	3871	2	N71302	HSV-1 gB and surround	3.13e-02
27	28	1.4	114	12	Q70473	Generic DNA sequence	2.27e-01
28	28	1.4	1454	1	N82107	PAP-I CDNA from human	2.27e-01
29	28	1.4	1466	2	N91353	Vascular anti-coagula	2.27e-01
30	28	1.4	1466	1	N80801	Sequence encoding vas	2.27e-01
31	28	1.4	1567	1	N81113	Plasmid contg. placen	2.27e-01
32	28	1.4	1575	1	N90112	Anticoagulant pp4 DNA	2.27e-01
33	28	1.4	1605	1	N91821	Endonexin II complete	2.27e-01
34	29	1.4	1840	11	Q65611	Feline zona pellucida	8.48e-02
35	26	1.3	31	17	Q99581	Human TPO anti-sense	1.54e+00
36	26	1.3	66	21	T13585	TSAR-9 library genera	1.54e+00
37	27	1.3	67	24	T14322	Primer used in the la	5.96e-01
38	27	1.3	82	21	T13610	DC43 TSAR library gen	5.96e-01
39	26	1.3	290	20	T21481	Human gene signature	1.54e+00
40	27	1.3	789	2	Q10792	Bovine vascular endot	5.96e-01
41	27	1.3	789	7	Q44255	Partial bovine VEGF-1	5.96e-01
42	27	1.3	829	2	Q10796	Bovine vascular endot	5.96e-01
43	27	1.3	961	2	Q10791	Bovine vascular endot	5.96e-01
44	27	1.3	961	7	Q44259	Bovine VEGF-164 codin	5.96e-01
45	26	1.3	1460	2	Q12679	PAP-I.	1.54e+00

ALIGNMENTS

RESULT 1
ID T23535 standard; cDNA to mRNA; 204 BP.
AC T23535;
DT 02-SEP-1996 (first entry)
DE Human gene signature HUMGS05382.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PE 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1393; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 204 BP; 54 A; 59 C; 35 G; 55 T;

QY 1562 GCCCAAACTTGTTGGGTCACAAACCCCTGTTGAGATAAAGCTGGCTGTATCTCAACATCTT 1621

Db 1015 nnnnnrgdvnmkgkgrmtywnngnrk 1041

QY 1622 CATCAGCTCCAGACTGAGACTCAGTGT 1648

RESULT 3

ID Q10572 standard; DNA; 1047 BP.

AC Q10572;

DT 09-APR-1991 (first entry)

DE Human Natriuretic Peptide Receptor B.

KW NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;

KW hyperaldosteronism; glaucoma; guanyl cyclase.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Peptide 1..22

FT /label= signal sequence

FT Protein 12

FT /label= mature NPBR

FT Domain 23..455

FT /label= extracellular domain

FT /note= "binds natriuretic peptides A,B and C]"

FT Domain 456..456

FT /label= transmembrane domain

FT Domain 479..1047

FT /label= cytoplasmic domain

FT /note= "GC and protien kinase activity"

FT Modified -site 24..26

FT /label= N-glycos_site

FT Modified -site 35..37

FT /label= N-glycos_site

FT Modified -site 161..163

FT /label= N-glycos_site

FT Modified -site 195..197

FT /label= N-glycos_site

FT Modified -site 244..246

FT /label= N-glycos_site

FT Modified -site 277..279

FT /label= N-glycos_site

FT Modified -site 349..351

FT /label= N-glycos_site

FT Modified -site 600..602

FT /label= N-glycos_site

PN WO9100292-A.

PD 10-JAN-1991.

PF 22-JUN-1990; U03586.

PR 23-JUN-1989; US-370673.

PA (GETH) GENENTECH INC.

PI Chang M, Goeddel D, Lowe D;

DR WPI; 91-036711/05.

DR N-PSDB; Q10324.

PT Natriuretic protein receptor B - for diagnosis and treatment of

PT kidney failure, heart failure, hyperaldosteronism, glaucoma etc.

PS Claim 3; Fig 1; 49pp; English.

CC The sequence was derived from the DNA encoding natriuretic peptide

CC receptor B, NPRB, having guanyl cyclase (GC) activity and protein

CC kinase activity. The DNA can be inserted into expression vectors

CC for the prodn. of the protein, opt. after being mutated to produce

CC NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr=

CC 114,952). The protein (or variants) can be used in treatment of

CC natriuretic peptide disorders, and also to isolate peptides using

CC affinity chromatography. Antibodies with affinity for NPRB can

CC also be prepd.

SQ Sequence 1047 BP; 87 A; 15 C; 83 G; 51 T;

Query Match 4.1%; Score 84; DB 2; Length 1047;

Best Local Similarity 9.1%; Pred. No. 5.69e-30;

Matches 93; Conservative 286; Mismatches 627; Indels 12; Gaps 11;

Db 28 avnnnnhnsyawaavrvgnavanavnangrannvdrnvssnnngacsnnannnsavdn 87

Cp 1222 AACGGGAGCTGAGATCCCGGAGCAGAAAATGGTCAGCCGTGCTCTGGAGCAGGCCTGTCTCG 1163

Db 88 knyhdndnnngcvcynaasvarnashwrnnnnntagavasgn-sakndhytrnvtgnsa 146

Cp 1162 AGCCATCAGGATGCCGGGACTAGCACCTGCAGCTTGGGATGCTGCGGGAGATGCGCTTGA 1103

Db 147 nkngnnvvtbhghnnwtaraannnyndartddrnhyntnngvnnanngsnsvnhnyarn 206

Cp 1102 ACTCTCTGCCCCCTTCTGCCACCGCTTCACGAGGTGATCAC-CAGCTCCCCACCCCTGT 1044

Db 207 nggnnnathnrangrnvycgnnnmnhnnnnanrnnntngdyvnnynvngnsnrag 266

Cp 1043 TTCTGTCCCATGACCAGATAGGGCGGTTGATGTCGTTTCATCTCCTCACAGGTGCACCTGC 984

Db 267 tratgrnwdrtrnnanan-rnanntvntyrnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn 325

Cp 983 AAGCTGTCTTTGAGCCACACGACCCGATTTCTTCAGGTCCTTTTCGGACACACCGTTTCAGC 924

Db 326 gnsnmnnnagcnydgnnnyanvnnntnnnggrndgrnrvnkmgrryhgvtgnvmdkn 385

Cp 923 TTGTAATGGTCTTGCTCTTGGTCTCCAGGATGATTT-TGGTATCT-CGGTGTGATGTAGG 866

Db 386 ndrntdvnwamgdndsgdnnaahysganknwwtgrnnnwkvganndhnnccanddd 445

Cp 865 TTATCTCCTTCACTTTTATTTCAGTGCAAAATCATT-TTTACAAAGCGTTTCCATTATG 807

Db 446 nscdktnnstnvanvgtnnnmngvssnnnnrknmnnknnsnmwrrnwnnnngnsn 505

Cp 806 TCGTGTGTCATCATCATTTT-TTATTTTTCAGGCTTCACAT--ACCTTTGGAGCTTCCTCG 750

Db 506 ryhkgagsrntnsnrgssygsnmtahgkynnnantghnknvvanhkvkkrnnntnrvn 565

Cp 749 GT-GGCTGGCAGAGGTGGTGGCTGCTAGCGAGGGGGATGCAAAAGGTGCTGTCTCTGGGG 691

Db 566 nnnkhmrdvnnhnrnngaacndnnnnncvtnrcnrgsnndnnnnndnnnnnnnnnn 625

Cp 690 GAAACGGTCCGACTCAAGCATGTCCGGGCCAGGGGAAGCCGAAGCGGACATGACCGGGGC 631

Db 626 dnvkgmannhnsnsshgskssncvdsrnvnkntdygnasnrstannndnnanyakkn 685

Cp 630 GCAGCGGTCTTCCCTGCACCTGCAGCAGAGCGAGTGGCATGGTGGTCTCGTCTAGGTC 571

Db 686 ntannnns-gnnnttgmaadvysngnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn 744

Cp 570 ATCGAGGCAGACGGGGGGCGAAGAGCGGACGACAGGAACCTTCTTGGTGTCCGGGTGGCAGT 511

Db 745 rnyrnrsndrtnnnnnnnnnnmrncwandhanrndngnnknrrnnknnggtsnndnnnr 804

Cp 510 CTTTCATGACCAGCGGGGATCCAAGCGCGCGCTGCTCCAGCACCTCCTTCATGGTCTCGTG 451

Db 805 mnnnyannnnknvnnrtnaynnnnkrkanannnnnnnnnnhsvannnnkrgntvnanandsvtyn 864

Cp 450 GCCCAGCAGGTGGGCAGCCCGCATGTTCTGTGTTTCGATGCGCGTGGCACAGCTGCAGGTT 391

Db 865 sdvngntansanstnmvvtnnndnytcndannndndvykvtngdaymvsngnngnrgn 924

Cp 390 GGCCGGGATGGGCTTGCAATTGCTGCGCTGTAGGAGAAAGTC-GGGCTGGCCAAAGAGGA 332

Db 925 rhannnarnananndavssnrnrhrnhdnrrnrgvhtgnvcagvvgnkmrncnngdtv 984

Cp 331 AGAGCCCCGCCGCCGAGCCAGGACAGTGGCAGGCGAGGAGGAGGAGCAGCAGCAGCAGC 272

Db 985 ntasrmnsnknanknhvssttkdandhngcnnnnnnnrgrdvnmkgkgrmtywnngnrkg 1042

Cp 271 CAGGGCCCTTCAGCATCGTGGCGCGGACCCCGAGGGGGCGAGAGGGGAGCGGACCGCG 214

RESULT 4

ID Q51746 standard; cDNA; 91 BP.

AC Q51746;

DT 31-MAY-1994 (first entry)

DE Oligonucleotide probe MK14-A

KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;

KW ss.

KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR P-PSDB; R65153.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70467 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)16(TGC)(NNB)12(NNB)16(TGC)(NNB)1Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are concatenated heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compsns. comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radioisotope, peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need for
CC complex methods of hybridoma formation or in vivo antibody production.
CC The TSARs are easily characterised and have designed activity allowing
CC direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 1.9%; Score 38; DB 12; Length 114;
Best Local Similarity 4.5%; Pred. No. 7.00e-06;
Matches 5; Conservative 34; Mismatches 73; Indels 0; Gaps 0;

Db 3 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 62
QY 198 GCGGCTCGCCCTTCCCGGCTCCGCTCCGCTCCCTCCCTCCGCTCGGGTCCGCCCGCCACGAT 257

Db 63 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 114
QY 258 GCTGCAGGGCGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 309

RESULT 13
ID Q70469 standard; DNA; 114 BP.
AC Q70469;
DT 07-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a

FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
DR Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure; Page 35; 255pp; English.
CC Q70469 is a generic DNA sequence used to generate random TSAR peptide
CC This generic formula can be represented as follows: X(TGC)(NNB)10-
CC (TGC)(NNB)6Z(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
CC sites (X is not the same as Y) that are not specified further. This
CC sequence generates peptides that are cloverleaf in structure. Other
CC generic sequences are shown in Q70465-68. Other specific peptides
CC generated by these generic sequences are shown in R65150-54. TSARs are
CC concatenated heterofunctional proteins or peptides, comprising at least
CC two functional regions - a binding domain with affinity for a ligand and
CC a second effector peptide portion that is chemically or biologically
CC active. They may further comprise a linker peptide between the 2 domains.
CC The oligonucleotides are also designed so that the expressed peptide
CC contains 2 or 4 cysteine residues positioned in, or flanking, the
CC unpredicted or variant residues. These residues confer some degree of
CC conformational rigidity to the peptides. The TSARs or compsns. comprising
CC a TSAR binding domain can be used in vivo to deliver a chemically or
CC biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC or enzyme, to the specific target or on the cell. They can also replace
CC the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC and therefore circumvent the need for complex methods of hybridoma
CC formation or in vivo antibody production. The TSARs are easily
CC characterised and have designed activity allowing direct and rapid
CC detection in a screening process.
SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

Query Match 1.9%; Score 39; DB 12; Length 114;
Best Local Similarity 7.3%; Pred. No. 2.34e-06;
Matches 8; Conservative 32; Mismatches 69; Indels 0; Gaps 0;

Db 6 bnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnnn 65
Cp 540 CAGGAACCTTCTTGGTGTCGGGTGGCACTGCTTCATGACGAGGGATCCAAGCGCGGC 481

Db 66 btgcnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnbnnb 114
Cp 480 CTGCTCCAGCACCTCCCTCATGGTCTCGTGGCCCCCAGCAGGTTGGGCAGC 432

RESULT 14
ID Q70466 standard; DNA; 114 BP.
AC Q70466;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR-9 petide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag= a
FT /note= "this sequence represents 'Z'; Z can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

Mpsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 9 09:11:05 1997; MasPar time 470.39 Seconds
965.771 Million cell updates/sec

Tabular output not generated.

Title: >848439-2-trans
Description: (1-885) from translate.seq
Perfect Score: 734
N.A. Sequence: 1 ATGYTNCARGNCCNGGNS.....SNATNMGNAARYTNCARTGY 885
Comp: TACRANGTYCCNGGCCNSW.....WNTANKCNTTYRANGTYACR

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 707517 seqs, 256659390 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:	EST-STS	
	1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8 9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14 15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20 21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26 27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32 33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38 39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44 45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50 51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56 57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62 63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68 69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74 75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80 81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86 87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92 93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98 99:EST99	
Database:	EST-STS-TWO	
	100:EST100 101:EST101 102:EST102 103:EST103 104:EST104 105:EST105 106:EST106 107:EST107 108:EST108 109:EST109 110:EST110 111:EST111 112:EST112 113:EST113 114:EST114 115:EST115 116:EST116 117:EST117 118:EST118 119:EST119 120:EST120 121:EST121 122:EST122 123:EST123 124:EST124 125:EST125 126:EST126 127:EST127 128:EST128 129:EST129 130:EST130 131:EST131 132:EST132 133:EST133 134:EST134 135:EST135 136:EST136 137:EST137 138:EST138 139:EST139 140:EST140 141:EST141 142:EST142 143:EST143 144:EST144 145:EST145 146:EST146 147:EST147 148:EST148 149:EST149 150:EST150 151:EST151 152:EST152 153:EST153 154:EST154 155:EST155 156:EST156 157:EST157 158:EST158 159:EST159 160:EST160 161:EST161 162:EST162 163:EST163 164:EST164 165:EST165 166:EST166 167:EST167 168:EST168 169:EST169 170:EST170 171:EST171 172:EST172 173:EST173 174:EST174	

175:EST175 176:EST176 177:EST177 178:EST178 179:EST179
180:EST180 181:EST181 182:EST182 183:EST183 184:EST184
185:EST185 186:EST186 187:EST187 188:EST188 189:EST189
190:EST190 191:EST191 192:EST192 193:EST193 194:EST194
195:EST195 196:EST196 197:EST197 198:EST198

Statistics: Mean 15.226; Variance 4.441; scale 3.429

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	345	47.0	462	174	W77968	zd70e11.r1 Soares fet	0.00e+00
2	319	43.5	396	178	AA073862	mm97f07.r1 Stratagene	0.00e+00
3	304	41.4	426	72	H87071	ys74d07.r1 Homo sapie	0.00e+00
4	279	38.0	406	61	H16121	yl28c05.r1 Homo sapie	0.00e+00
5	277	37.7	414	156	AA024771	ze76g10.r1 Soares fet	0.00e+00
6	242	33.0	344	61	H14917	yl26d03.r1 Homo sapie	5.47e-272
7	210	28.6	299	148	W08345	mb41f02.r1 Soares mou	1.74e-227
8	163	22.2	231	65	H29095	ym31e03.r1 Homo sapie	4.01e-163
9	151	20.6	250	172	AA061047	mj78a12.r1 Soares mou	5.97e-147
10	63	8.6	531	12	M89402	CEL08B2 Caenorhabditi	2.81e-35
11	57	7.8	360	141	C09584	C.elegans cDNA clone	1.49e-28
12	52	7.1	360	89	CELK117B4F	C.elegans cDNA clone	3.90e-23
13	41	5.6	360	89	CELK079D1F	C.elegans cDNA clone	5.43e-12
14	40	5.4	256	47	H45312	yo65b03.r1 Homo sapie	4.79e-11
15	37	5.0	206	101	N63268	yz88g01.s1 Homo sapie	2.72e-08
16	37	5.0	418	101	N62714	yz76d05.s1 Homo sapie	2.72e-08
17	36	4.9	341	77	CELK119F4R	C.elegans cDNA clone	2.10e-07
18	34	4.6	245	129	HSC0NC092	H. sapiens partial cd	1.11e-05
19	34	4.6	336	123	HSBA0B082	H. sapiens partial cd	1.11e-05
20	33	4.5	227	184	AA017479	ze38h10.r1 Soares ret	7.63e-05
21	33	4.5	350	187	AA144540	mr17f12.r1 Soares mou	7.63e-05
22	33	4.5	414	194	AA169546	zo89b08.r1 Stratagene	7.63e-05
23	33	4.5	428	95	N41923	yy07a02.r1 Homo sapie	7.63e-05
24	33	4.5	433	194	AA169521	zo89a08.r1 Stratagene	7.63e-05
25	33	4.5	443	113	W03474	za08a10.r1 Soares mel	7.63e-05
26	33	4.5	458	23	R42865	yg06a10.s1 Homo sapie	7.63e-05
27	32	4.4	219	129	HSC0EA112	H. sapiens partial cd	4.99e-04
28	32	4.4	226	156	AA027879	zk05e04.r1 Soares pre	4.99e-04
29	32	4.4	242	106	HSC29E102	H. sapiens partial cd	4.99e-04
30	32	4.4	242	119	W44804	zb98g09.s1 Soares par	4.99e-04
31	32	4.4	243	133	N84441	KK8808F Homo sapiens	4.99e-04
32	32	4.4	253	135	R58334	G2318 Fetal heart Hom	4.99e-04
33	32	4.4	262	81	T35329	EST82912 Homo sapiens	4.99e-04
34	32	4.4	345	105	HSC1EE122	H. sapiens partial cd	4.99e-04
35	32	4.4	384	95	N42176	yy07e11.r1 Homo sapie	4.99e-04
36	32	4.4	453	95	N42436	yy09h05.r1 Homo sapie	4.99e-04
37	32	4.4	473	127	W90442	zh78c02.s1 Soares fet	4.99e-04
38	32	4.4	584	4	T67709	yc27c11.s1 Homo sapie	4.99e-04
39	31	4.2	252	109	HUM069H03B	Human fetal brain cDN	3.11e-03
40	31	4.2	293	22	R38715	yd03b08.s1 Homo sapie	3.11e-03
41	31	4.2	295	26	R52481	yg80c12.s1 Homo sapie	3.11e-03
42	31	4.2	362	24	R46624	yg53c02.s1 Homo sapie	3.11e-03
43	31	4.2	372	72	H87471	yw17f06.s1 Homo sapie	3.11e-03
44	31	4.2	390	1	CELK025G2R	C.elegans cDNA clone	3.11e-03
45	31	4.2	483	98	N52277	yv47e12.s1 Homo sapie	3.11e-03

ALIGNMENTS

RESULT	1	W77968	462 bp	mrna	EST	17-OCT-1996
LOCUS	zd70e11.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone					
DEFINITION	346028 5'.					
ACCESSION	W77968					
NID	91388502					
KEYWORDS	EST.					


```
QY 576 YGAYTTYGCNYTNAARATNAARGTNAARGARATNACNTAYATNAAYMGNGAYAC 629

RESULT 8
LOCUS H29095 231 bp mRNA EST 17-JUL-1995
DEFINITION Ym31e03.r1 Homo sapiens cDNA clone 49614 5'.
ACCESSION H29095
NID g900005
KEYWORDS EST.
SOURCE human clone=49614 library=Soares infant brain INIB vector=Lafmid BA
host=DH10B (ampicillin resistant) primer=M13RP1 Rsite1=Not I
Rsite2=Hind III whole brain from a 73 days post natal female. 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
AACTGGAAGAATTCGGCGCCGACGAATTTTCTTTTCTTTTCTTTT 3']; double-stranded
cDNA was ligated to Hind III adaptors (Pharmacia), digested with
Not I and directionally cloned into the Not I and Hind III sites of
the Lafmid BA vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.Fatima
Bonaldo.

ORGANISM Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 231)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT
GDB: G00-422-155
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 127
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES
source
1..231
/organism="Homo sapiens"
/clone="49614"

BASE COUNT 70 a 54 c 58 g 39 t 10 others
ORIGIN
Query Match 22.2%; Score 163; DB 65; Length 231;
Best Local Similarity 55.2%; Pred. No. 4.01e-163;
Matches 128; Conservative 51; Mismatches 51; Indels 2; Gaps 2;

Db 1 cactgaaaaataaantnaaggagataacctaaccacccgagataccaaannccatcctgg 60
| : ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : |||
QY 584 CNYTNAARATNAARGTNAARGARATNACNTAYATNAAYMGNGAYACNAARATNATNTYNG 643

Db 61 gagaccaagagcaagaccatttacagctgaacgggtgtgtccganag-gacctgaagaat 119
| : ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : |||
QY 644 -ARACNAARWSNAARACNATNTAYAARYTNAAYGGNGTNSNGARNGNGAYYTNAARAAR 702

Db 120 tcggtcctgtggctcanagacagcttgcaagtgcacctgttnaggagatgaacgacatcaac 179
: : ||| ||| : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
QY 703 WSNGTNTYNTGGYTNAARGAYWSNYTNCARTGYACNTGYGARGARATGAAYATNAAY 762

Db 180 gcgccttatctggncatgggacagacacaggggtngggagctgggtgntcacct 231
|| || ||| : ||| ||| ||| : ||| ||| ||| : ||| ||| ||| : ||| :
QY 763 GCNCCNTAYTYTNGTNGTATGGGNCARARCARCGGNGGNGARYTNGTNGTATNACNW 814
```

```
RESULT 9
LOCUS AA061047 250 bp mRNA EST 23-SEP-1996
DEFINITION mj78a12.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 482206
5'.
ACCESSION AA061047
NID g1554620
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.
REFERENCE 1 (bases 1 to 250)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:292950
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 179.
Location/Qualifiers
1..250
/organism="Mus musculus"
/note="Vector: pT7T3D (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTACCAATCTGAGTGGGAGCGCGCATTTTCTTTTCTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Fatima Bonaldo. RNA was kindly provided by
Dr. Minoru Ko (Wayne State University)."
```

/clone="482206"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
<1..>250
mRNA
BASE COUNT 55 a 71 c 75 g 49 t
ORIGIN
Query Match 20.6%; Score 151; DB 172; Length 250;
Best Local Similarity 56.1%; Pred. No. 5.97e-147;
Matches 105; Conservative 48; Mismatches 34; Indels 0; Gaps 0;

Db 1 gaaatccgtgctgtggtcctcaaaagacagcctgcagtcgcctgtgaggagatgaacgacat 60
: ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 699 RAARWSNGTNYTNTGGYTNAARGAYWSNYTNCARTGYACNTGYGARGARATGAAYGAYAT 758

Db 61 caacgctccgtatctgtgtcatctgggcagacagagcaggcgcgagctggtgatgacctcgt 120
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : |||
QY 759 NAAYGCNCCTAYTYTNGTNGTATGGGNCARARCARCGGNGGNGARYTNGTNGTATNACNWSNGT 818

Db 121 gaaacggtggcagaaggccagagatagttcaagcgcatctcccgcagcatccgcaagct 180
||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 819 NAARMGNTGGCARAARGGNCARMGNGARTTYAARMGNATNWSNMGNWSNATNMGNAARYT 878

Db 181 gcaatgc 187
||| :

QY 879 NCARTGY 885

RESULT 10

LOCUS M89402 531 bp mRNA EST 02-DEC-1992

DEFINITION CEL08B2 Caenorhabditis elegans cDNA clone cm08b2 5' similar to gene

frizzled protein precursor - Drosophila homologous peptide.

ACCESSION M89402

NID g275297

KEYWORDS EST.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Waterston R.H.(USA) and Sulston J.E.(UK)

(USA) Dept. of Genetics or (UK)

(USA) Washington Univ. School of Medicine or (UK) MRC Laboratory of

Molecular Biology

Box 8232,4566 Scott Ave.,St. Louis,MI 63110,USA,or,Hills

Road,Cambridge CB2 2QH,UK

Tel: (USA) (314)3627072 or (UK) (0223)248011

Fax: (USA) (314)3624137 or (UK) (0223)402008

Email: rw@nematode.wustl.edu or jes@nrc-lmb.cam.ac.uk.

Location/Qualifiers

1..531

/organism="Caenorhabditis elegans"

/clone="cm08b2"

/strain="Bristol N2"

BASE COUNT 130 a 123 c 136 g 138 t 4 others

ORIGIN

Query Match 8.6%; Score 63; DB 12; Length 531;

Best Local Similarity 37.5%; Pred. No. 2.81e-35;

Matches 118; Conservative 46; Mismatches 151; Indels 0; Gaps 0;

Db 117 atcatgagatgtgcaacgatttgccgtataacttaacgagcttcccaaatctcgtcgacg 176

||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|

QY 137 AYTNCARYTNTGYCAYGGNATNGARTAYCARAAAYATGMGNYTNCNAAAYTNTYNGNC 196

Db 177 aggaatcatggaagagcctccgaatccatcctccacctaagaagccctgctcctcggtg 236

||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|

QY 197 AYGARACNATGAARGARGTNYTNGARCARGCNGGNGCCTGGATNCNCTNYTNGTNGTGAARC 256

Db 237 tctgctccgagcagctcaaatcttccctgtgctccgtctacttcccgatgtgcaacagaga 296

||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|

QY 257 ARTGYCAYCCNGAYACNAARAARTTYTNTGYWSNTYNTTYGNCNCCNGTNTGYTNGAYG 316

Db 297 aactagccaacccaattggtcccatgccgtccattgtgttcttccgtccaggaaaagtgtc 356

||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|

QY 317 AYTNGAYGARACNATNCARCCNTGYCAYWSNTYNTGYTNCARGTNAARGAYMGNTGYG 376

Db 357 ttccagtgctggaaagtctcggtttcaagtgggcccgatgtgattcgttgataagtctc 416

||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|

QY 377 CNCCNGTNGTWSNGCNTTYGGNTTYCCNTGGCCNGATGYTNGARTGYGAYMGNTTYC 436

Db 417 cgttngagagaacaatc 431

||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|

QY 437 CNCARGAAYAYGAY 451

RESULT 11

LOCUS C09584 360 bp mRNA EST 09-SEP-1996

DEFINITION C.elegans cDNA clone yk173b2 : 5' end, single read.

ACCESSION C09584

NID g1534655

KEYWORDS EST(expressed sequence tag).

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

Direct Submission

Submitted (29-APR-1996) to the DBJ/EMBL/GenBank databases. Yuji

Kohara, National Institute of Genetics, Gene Library Lab; Yata

1111, Mishima, Shizuoka 411, Japan (E-mail:ykohara@lab.nig.ac.jp,

Tel:0559-81-6854, Fax:0559-81-6855)

2 (sites)

Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A.,

Sano,M., Miyata,A. and Nishigaki,A.

Expression map of the C.elegans genome

Unpublished (1996)

Location/Qualifiers

1..360

/organism="Caenorhabditis elegans"

/strain="N2"

/clone="yk173b2"

/clone_lib="Yuji Kohara unpublished cDNA"

/dev_stage="embryo"

/sex="hermaphrodite"

BASE COUNT 86 a 90 c 90 g 87 t 7 others

ORIGIN

Query Match 7.8%; Score 57; DB 141; Length 360;

Best Local Similarity 41.9%; Pred. No. 1.49e-28;

Matches 72; Conservative 29; Mismatches 71; Indels 0; Gaps 0;

Db 181 ttctgtgctccgtctacttcccgatgtgcaacgagaaactagccaacccaattggtcca 240

||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|

QY 280 TYYTNTGYWSNTYNTTYGCNCNGTNTGYTNGAYGAYTNGAYGARACNATNCARCCN 339

Db 241 tgccgtccattgtgtcttccgtccaggaaaagtgtcttccagtgctggaaaagtctcgtt 300

||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|

QY 340 TGICAYWSNTYNTGYGTNCARGTNAARGAYMGNTGYGCNCNGTNGTNGWSNGCNTTYG 399

Db 301 ttcaagtgggcccgatgtgattcgggtgtgataaagttcccggttgagagaacaatc 352

||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|

QY 400 TTYCNTGGCCNGAYATGYTNGARTGYGAYMGNTTYCCNCARGAAYAYGAY 451

RESULT 12

LOCUS CELK117B4F 360 bp mRNA EST 13-DEC-1995

DEFINITION C.elegans cDNA clone yk117b4 : 5' end, single read.

ACCESSION D76340

NID g1122125

KEYWORDS EST(expressed sequence tag).

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (23-Aug-1995) to DDBJ by:

Yuji Kohara

Gene Library Lab.

National Institute of Genetics

Query Match 5.4%; Score 40; DB 47; Length 256;
Best Local Similarity 49.4%; Pred. NO. 4.79e-11;
Matches 39; Conservative 15; Mismatches 25; Indels 0; Gaps 0;

Db 134 agacgtccccctcggaancttgtcacacttaagcatctcgggccagtagaagcgaaga 193
::||| ||| ||| :||:|||| |:|||| ||||| :|| ||:|
Cp 451 RRTCTRTTCYTGNGRAANCKRTCRCAAYTCNARCATRTCTNGGCCANGGRAANCCRAANG 392

Db 194 actgcangaccggtcgca 212
:|| | ||| |:
Cp 391 CNWSCATNACNGGNCRCa 373

RESULT 15
LOCUS N63268 206 bp mRNA EST 01-MAR-1996
DEFINITION yz88d01.s1 Homo sapiens cDNA clone 290113 3'.
ACCESSION N63268
NID 91211097
KEYWORDS EST.
SOURCE human clone=290113 primer=ml3 -40 forward library=Soares multiple sclerosis 2NBHSP vector=pt7T3D (Pharmacia) with a modified polylinker V-type; phagemid host=DH10B (ampicillin resistant) Rsitel=Not I Rsite2=Eco RI 46 year old male. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGAGCGCGCATTTTTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from 4 multiple sclerosis lesions from one patient was kindly provided by Dr. Kevin G. Becker (NINDS/NIH).

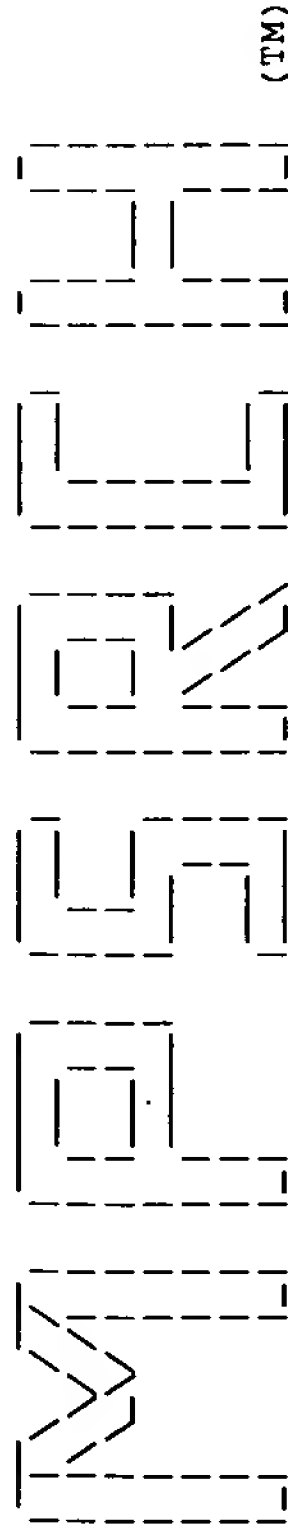
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygil; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 206)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.

TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 161
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES Location/Qualifiers
source 1..206
/organism="Homo sapiens"
/clone="290113"
mrna <1..>206

BASE COUNT 67 a 39 c 53 g 46 t 1 others
ORIGIN
Query Match 5.0%; Score 37; DB 101; Length 206;
Best Local Similarity 43.4%; Pred. NO. 2.72e-08;
Matches 36; Conservative 16; Mismatches 31; Indels 0; Gaps 0;

Db 5 aaataagagtggcagcataaggaatactatttataaaaaaacagagttatagggcta 64
|:|| | | | : | || | || || ||:| | | | | | : ||:
QY , 632 ARATNATNYTNGARACNAARWSNAARACNATNTAYAARYTNAAYGGNGTNSNGARMGNG 691

Db 65 ctttaaagaagaatgaactttgg 87
:| ||:| |: | | | |
QY 692 AYYTNAARAARWSNGTYNTTGG 714
Search completed: Tue Dec 9 09:33:30 1997
Job time : 1345 secs.



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

Mpsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Tue Dec 9 09:33:55 1997; MasPar time 244.79 Seconds
993.543 Million cell updates/sec
Tabular output not generated.

Title: >848439-2-trans
Description: (1-885) from translate.seq
Perfect Score: 734
N.A. Sequence: 1 ATGYTNCARGGNCNGGNWS.....SNATMGNAARYTNCARTGY 885
Comp: TACRANGTYCCNGGCCNSW.....WNTANKCNTTYRANGTYACR

Scoring table: TABLE default
Gap 6
Nmatch STD : Dbase 0; Query 0
Searched: 359085 seqs, 137405154 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: EST-STS-THREE
1:EST199 2:EST200 3:EST201 4:EST202 5:EST203 6:EST204
7:EST205 8:EST206 9:EST207 10:EST208 11:EST209 12:EST210
13:EST211 14:EST212 15:EST213 16:EST214 17:EST215
18:EST216 19:EST217 20:EST218 21:EST219 22:EST220
23:EST221 24:EST222 25:EST223 26:EST224 27:EST225
28:EST226 29:EST227 30:EST228 31:EST229 32:EST230
33:EST231 34:EST232 35:EST233 36:STS1 37:STS2 38:STS3
39:STS4 40:STS5 41:STS6 42:STS7 43:STS8 44:STS9 45:STS10
46:STS11 47:STS12 48:STS13
EST-STS-FOUR
49:gNEST1 50:gNEST2 51:gNEST3 52:gNEST4 53:gNEST5
54:gNEST6 55:gNEST7 56:gNEST8 57:gNEST9 58:gNEST10
59:gNEST11 60:gNEST12 61:gNEST13 62:gNEST14 63:gNEST15
64:gNEST16 65:gNEST17 66:gNEST18 67:gNEST19 68:gNEST20
69:gNEST21 70:gNEST22 71:gNEST23 72:gNEST24 73:gNEST25
74:gNSTs 75:enEST1 76:enEST2 77:enEST3 78:enEST4
79:enEST5 80:enEST6 81:enEST7 82:enEST8 83:enEST9
84:enEST10 85:enEST11 86:enEST12 87:enEST13 88:enEST14
89:enEST15 90:enEST16 91:enEST17 92:enEST18 93:enEST19
94:enEST20 95:enEST21 96:enEST22 97:enEST23 98:enEST24
99:enEST25 100:enEST26 101:enEST27 102:enEST28
103:enEST29 104:enEST30 105:enEST31 106:enSTS 107:ueEST1
108:ueEST2

Database: EST-STS-FOUR
49:gNEST1 50:gNEST2 51:gNEST3 52:gNEST4 53:gNEST5
54:gNEST6 55:gNEST7 56:gNEST8 57:gNEST9 58:gNEST10
59:gNEST11 60:gNEST12 61:gNEST13 62:gNEST14 63:gNEST15
64:gNEST16 65:gNEST17 66:gNEST18 67:gNEST19 68:gNEST20
69:gNEST21 70:gNEST22 71:gNEST23 72:gNEST24 73:gNEST25
74:gNSTs 75:enEST1 76:enEST2 77:enEST3 78:enEST4
79:enEST5 80:enEST6 81:enEST7 82:enEST8 83:enEST9
84:enEST10 85:enEST11 86:enEST12 87:enEST13 88:enEST14
89:enEST15 90:enEST16 91:enEST17 92:enEST18 93:enEST19
94:enEST20 95:enEST21 96:enEST22 97:enEST23 98:enEST24
99:enEST25 100:enEST26 101:enEST27 102:enEST28
103:enEST29 104:enEST30 105:enEST31 106:enSTS 107:ueEST1
108:ueEST2
Statistics: Mean 15.318; Variance 4.699; scale 3.259
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result § Query
SUMMARIES

No.	Score	Match	Length	DB	ID	Description	Pred.	No.
1	357	48.6	458	33	AA105749	m184a03.r1 Stratagene	0.00e+00	
2	319	43.5	396	54	AA073862	mm97f07.r1 Stratagene	0.00e+00	
3	235	32.0	394	101	MMAA60087	va52g08.r1 Soares mou	4.81e-249	
4	210	28.6	299	94	NM3451	mb41f02.r1 Soares mou	4.84e-216	
5	148	20.2	184	20	AA209833	mo90h01.r1 Beddington	1.07e-135	
6	145	19.8	523	56	AA122822	mr03c11.r1 Soares mou	6.95e-132	
7	145	19.8	523	6	AA122822	mr03c11.r1 Soares mou	6.95e-132	
8	136	18.5	287	32	AA082155	ze88f06.r1 Soares fet	1.67e-120	
9	55	7.5	502	21	AA020088	mh49d08.r1 Soares mou	5.43e-25	
c 10	38	5.2	213	28	AA023595	mh80d03.r1 Soares mou	1.05e-08	
11	37	5.0	421	82	HS1147060	zr18a11.s1 Stratagene	7.52e-08	
12	37	5.0	421	63	AA227473	zr18a11.s1 Stratagene	7.52e-08	
13	36	4.9	512	29	AA044908	zf51e10.r1 Soares ret	5.22e-07	
c 14	35	4.8	432	29	AA030216	mh87g08.r1 Soares mou	3.49e-06	
c 15	35	4.8	519	39	G08492	human STS CHLC.ATA11D	3.49e-06	
c 16	34	4.6	288	36	DM57F11T	D. melanogaster STS d	2.25e-05	
c 17	34	4.6	496	10	AA013810	mh06e01.r1 Soares mou	2.25e-05	
c 18	33	4.5	241	21	AA019228	ze59c03.r1 Soares ret	1.39e-04	
c 19	33	4.5	252	37	G03067	human STS WI-5108.	1.39e-04	
c 20	33	4.5	350	56	AA144540	mr17f12.r1 Soares mou	1.39e-04	
c 21	32	4.4	345	44	G27227	human STS SHGC-32136.	8.23e-04	
c 22	32	4.4	552	12	AA180957	zp41c03.r1 Stratagene	8.23e-04	
c 23	31	4.2	337	10	AA009696	ze82h05.s1 Soares fet	4.66e-03	
c 24	31	4.2	407	89	HSAA62672	zs23f08.r1 Soares NBH	4.66e-03	
c 25	31	4.2	417	6	AA122184	zn28c08.s1 Stratagene	4.66e-03	
c 26	31	4.2	444	17	AA197092	zq10e04.r1 Stratagene	4.66e-03	
27	31	4.2	451	39	G09558	human STS CHLC.GCT8C0	4.66e-03	
c 28	31	4.2	491	25	N53936	yv59d10.s1 Soares fet	4.66e-03	
c 29	31	4.2	516	12	AA181149	zp69c10.s1 Stratagene	4.66e-03	
c 30	30	4.1	177	36	DM52H6T	D. melanogaster STS d	2.51e-02	
c 31	30	4.1	213	10	AA018016	mh45c07.r1 Soares mou	2.51e-02	
32	30	4.1	278	105	TB2692	T1343 MVAT4 bloodstre	2.51e-02	
c 33	30	4.1	305	21	AA019656	ze61f11.s1 Soares ret	2.51e-02	
34	30	4.1	339	43	G26024	human STS EST186690.	2.51e-02	
35	30	4.1	402	1	N97824	1353C3 czappFDD2.1, D	2.51e-02	
36	30	4.1	407	21	AA022630	ze71h06.s1 Soares fet	2.51e-02	
37	30	4.1	425	43	G24362	human STS WI-13950.	2.51e-02	
c 38	30	4.1	435	34	AA141629	CK02453.3prime Drosop	2.51e-02	
39	30	4.1	440	26	N64746	yz30a03.s1 Soares mul	2.51e-02	
40	30	4.1	455	62	AA224647	CpEST.119 psKIIminusC	2.51e-02	
41	30	4.1	455	80	CPAA24647	CpEST.119 psKIIminusC	2.51e-02	
c 42	30	4.1	471	44	G27881	human STS SHGC-34074.	2.51e-02	
43	30	4.1	503	70	AA241149	mv26g08.r1 GuayWoodfo	2.51e-02	
44	30	4.1	503	92	MM1157919	mv26g08.r1 GuayWoodfo	2.51e-02	
45	30	4.1	527	9	AA131693	zl34f05.s1 Soares pre	2.51e-02	

ALIGNMENTS

RESULT	1	AA105749	458 bp	mrna	EST	04-FEB-1997
LOCUS		m184a03.r1 Stratagene mouse kidney (#937315)				
DEFINITION		clone 518668 5' similar to TR:G1151260 G1151260 TRANSMEMBRANE RECEPTOR. ; .				
ACCESSION		AA105749				
NID		g1654838				
KEYWORDS		EST.				
SOURCE		house mouse.				
ORGANISM		Mus musculus				
		Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE		1 (bases 1 to 458)				
AUTHORS		Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.				
		The WashU-HMI Mouse EST Project				
TITLE		Unpublished (1996)				
JOURNAL						
COMMENT						


```
QY 521 CNTGYAARAAYAAAGAYGAYGAYAAATGATGACACNTGTAARAAYGAYT 580
Db 359 tcgcactgaaatcaaaagtgaaggagataaacgtacat 395
QY 581 TYGCNYTNAARATNAARGTNAARGARATNACNTAYAT 617

RESULT 3
ID MAA60087 standard; RNA; EST; 394 BP.
AC AA260087;
NI g1896590
DT 19-MAR-1997 (Rel. 51, Created)
DT 19-MAR-1997 (Rel. 51, Last updated, Version 1)
DE va52g08.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone 735038 5'.
KW EST.
OS Mus musculus (house mouse)
OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC Mus.
RN [1]
RP 1-394
RA Maria M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA Waterston R.;
RA "The WashU-HHMI Mouse EST Project";
RL Unpublished.
CC Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC Washington University School of Medicinep 4444 Forest Park Parkway,
CC Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC Email: mouseest@watson.wustl.edu This clone is available
CC royalty-free through LLNL; contact the IMAGE Consortium
CC (info@image.llnl.gov) for further information. MGI:452086 Seq
CC primer: -28ml3 rev2 ET from Amersham.
FH Key Location/Qualifiers
FH source
FH 1..394
FH /organism="Mus musculus"
FH /strain="C57BL/6J"
FH /note="Vector: pT7T3D-pac (Pharmacia) with a modified
FT polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
FT was primed with a Not I - oligo(dT) primer [5',
FT TGTTACCAATCTGAAGTGGGAGCGGCCCTATTATTTTTCCTTTT 3'], o
n
FT total mouse RNA [providedby Minoru Ko, Wayne State Univ.];
FT double-stranded cDNA was ligated to Eco RI adaptors
FT (Pharmacia), digested with Not I and cloned into the Not I
FT and Eco RI sites of the modified pT7T3 vector. Library wen
t
FT through one round of normalization, and was constructed by
FT Bento Soares and M. Fatima Bonaldo."
FT /clone="735038"
FT /clone_lib="Soares mouse 3NME12 5"
FT /sex="unknown"
FT /dev_stage="12.5dpc total fetus"
FT /lab_host="DH10B"
FT mRNA
FT Sequence 394 BP; 104 A; 96 C; 118 G; 76 T; 0 other;
SQ
Query Match 32.0%; Score 235; DB 101; Length 394;
Best Local Similarity 57.5%; Pred. No. 4.81e-249;
Matches 164; Conservative 71; Mismatches 50; Indels 0; Gaps 0;

Db 1 aaggagataacgtacatcaacagagacaccaaagatcatcctgagacaaagacacc 60
QY 601 AARGARATNACNTAYATNAAYMGNAYACNAARATNATNTNGARACNAARWSNAARACN 660
Db 61 atttacaagctgaacggcgtgtccgaaaggacctgaagaaatccgtgtggtctcaaa 120
QY 661 ATNTAYAARYTNAAYGGNGTNGARMGNNGAYTNAARAARWSNGTNTYNTGGYTNAAR 720
Db 121 gacagcctgcagtcacctgtgaggagatgaacgacatcaacgctccgtatctgtgctatg 180
```

```
QY 721 GAYWSNYTNCARTGYACNTGYGARGARATGAAYGAYATNAAYGCNCNTAYTNGTNATG 780
Db 181 ggacagaagcaggcgcgagctgtgtagtcacctccgtgaaacgggtggcagaaggccag 240
QY 781 GGNCARAARCARGGNGGNGARYTNGTNGTNATNACNWSNGTNAARMGNTGGCARAARGNCAR 840
Db 241 agagagttcaagcgcatctcccgcagcatccgcaagctgcaatgc 285
QY 841 MGNARTTYAARMGNATNWSNMGNWSNATNMGNMNAARYTNCARTGY 885

RESULT 4
ID MM3451 standard; RNA; EST; 299 BP.
AC W08345;
NI g1282366
DT 27-APR-1996 (Rel. 47, Created)
DT 05-MAR-1997 (Rel. 51, Last updated, Version 2)
DE mb41f02.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 331995
DE 5'.
KW EST.
OS Mus musculus (house mouse)
OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC Mus.
RN [1]
RP 1-299
RA Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA Waterston R.;
RA "The WashU-HHMI Mouse EST Project";
RL Unpublished.
CC Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC Washington University School of Medicinep 4444 Forest Park Parkway,
CC Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC Email: mouseest@watson.wustl.edu This clone is available
CC royalty-free through LLNL; contact the IMAGE Consortium
CC (info@image.llnl.gov) for further information. MGI:213395 Seq
CC primer: mob.REGA+ET High quality sequence stop: 280.
FH Key Location/Qualifiers
FH source
FH 1..299
FH /organism="Mus musculus"
FH /note="Vector: pT7T3D (Pharmacia) with a modified
FT polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
FT was primed with a Not I - oligo(dT) primer [5',
FT TGTTACCAATCTGAAGTGGGAGCGGCCCATTTTTCCTTTTTCCTTTT 3'],
FT double-stranded cDNA was size selected, ligated to Eco RI
FT adapters (Pharmacia), digested with Not I and cloned into
FT the Not I and Eco RI sites of a modified pT7T3 vector
FT (Pharmacia). Library went through one round of
FT normalization to a Cot = 5. Library constructed by Bento
FT Soares and M.Fatima Bonaldo. RNA was kindly provided by Dr
FT Minoru Ko (Wayne State University)."
FT /clone="331995"
FT /clone_lib="Soares mouse p3NMF19.5"
FT /dev_stage="19.5 dpc total fetus"
FT /lab_host="DH10B (ampicillin resistant)"
FT mRNA
FT Sequence 299 BP; 79 A; 95 C; 72 G; 53 T; 0 other;
SQ
Query Match 28.6%; Score 210; DB 94; Length 299;
Best Local Similarity 59.5%; Pred. No. 4.84e-216;
Matches 175; Conservative 68; Mismatches 46; Indels 5; Gaps 5;

Db 11 gcctcgtcactcgtctct-cgtgcag-tgaaggaccgctgcgc-cgggtcatgtccgcctt 67
QY 336 RCCNTGTCAYWSNYTNTGYGTNCARGTNAARGAYMGNATGYGCNCNGTNGTNGSNGCNTT 395
Db 68 cg-cttccctggccagacatgctggagtgcgaccgtttcccgag-acaacgacctctg 125
```

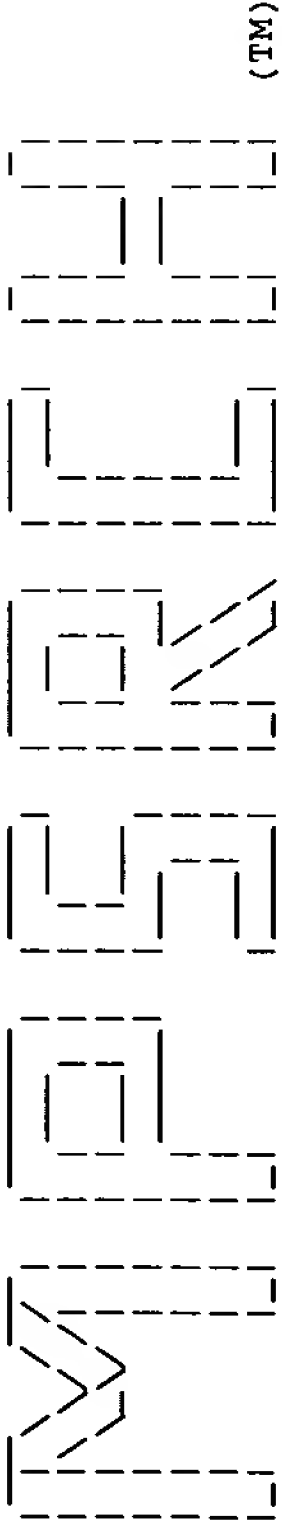
QY	396	YGGNTTYCCNTGGCCNGAYATGYTNGARTGYGAYMGNTTYCCNCARGAAYAAYGAYYNTTG	455
Db	126	catcccccctgctagtagcgaccacacctctgcgcggccacagaggaagctccccaaggtgtg	185
QY	456	YATNCCNTYNGCNWSNGAYCAYYTYNTNCCNGCNACNGARGARGCNCNAARGTNTG	515
Db	186	tgaagcctgcacaaaccaaagaatgaggacgacacacgacatcatcgaaacccctttgtaaaaa	245
QY	516	YGARGCNTGYAARAAYAARAAYGAYGAYAAAYGAYATNATGGARACNYTNTGYAARAA	575
Db	246	tgacttcgcactgaaaaatcaaaagtgaaagagataaacgtacatcaacacagagacac	299
QY	576	YGAYTTYGCNYTNAARATNAARGTNAARGARATNACNTAYATNAAYMGNAYAC	629
RESULT	5		
LOCUS	AA209833	184 bp	mRNA
DEFINITION	mo90h01.r1 Beddington mouse embryonic region Mus musculus cDNA	EST	29-JAN-1997
ACCESSION	AA209833		
NID	g1808152		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;		
	Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Mus.		
REFERENCE	1 (bases 1 to 184)		
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Scheellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.		
TITLE	The WashU-HHMI Mouse EST Project		
JOURNAL	Unpublished (1996)		
COMMENT	Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project Washington University School of Medicinep 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:341673 Seq primer: -28ml3 rev1 ET from Amersham High quality sequence stop: 147. Location/Qualifiers 1..184 /organism="Mus musculus" /strain="C57BL6 x DBA" /note="Vector: pCMV-SPORT; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Gastrulating embryos were collected at 7.5dpc from C57BL6 x DBA matings, excluding embryos that had developed head folds and all extraembryonic tissues. Average insert size: 1.3 kb (range: 0.5 - 3.0 kb). Referenced in Development 121, 2479-2489 (1995)" /clone="560881" /clone_lib="Beddington mouse embryonic region" /sex="pooled" /dev_stage="7.5dpc" /lab_host="DH12S" <1..>184		
FEATURES	source		
	61 a	45 c	48 g 30 t
BASE COUNT	61 a	45 c	48 g 30 t
ORIGIN			
	Query Match	20.2%;	Score 148; DB 20; Length 184;
	Best Local Similarity	58.9%;	Pred. No. 1.07e-135;
	Matches	109; Conservative	45; Mismatches 30; Indels 1; Gaps 1;

Db	1	gataacgtacatcaacagagacaccaagatcatcctctggagacaaaagacaagaccattta	60
QY	606	RATNACNTAYATNAAAYMGNGAYACNAARATNATNTNGARACNAARWSNAARACNATNTA	665
Db	61	caagctgaacggcgtgtccgaaaag-gacctgaagaaaatccgctgctgtggtctcaagacag	119
QY	666	YAARYTNAAYGGNGTWSNGARMGNGAYYTNAARAARWSNGTNYTNTGGYTNAARGAYS	725
Db	120	cctgcagtcacctgtgaggagatgaacgacatcaacgctccgtatctctggtcatgggaca	179
QY	726	NYTNCARTGYACNTGYGARGARATGAAYGAYATNAAAYGCNCCNTAYYTNGTNATGGNCA	785
Db	180	gaagc 184	
QY	786	RAARC 790	
RESULT	6		
LOCUS	AA122822	523 bp	mRNA
DEFINITION	mr03c11.r1 Soares mouse 3NbMS Mus musculus cDNA clone 596372 5'.	EST	16-FEB-1997
ACCESSION	AA122822		
NID	g1681779		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;		
	Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Mus.		
REFERENCE	1 (bases 1 to 523)		
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Scheellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wllson,R. and Waterston,R.		
TITLE	The WashU-HHMI Mouse EST Project		
JOURNAL	Unpublished (1996)		
COMMENT	Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project Washington University School of Medicinep 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:361804 Seq primer: -28Ml3 rev2 from Amersham High quality sequence stop: 469. Location/Qualifiers 1..523 /organism="Mus musculus" /strain="C57BL/6J" /note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5, TGTTACCAATCTGAAGTGGGAGCGCGCGCTGTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Bertrand Jordan. Library went through three rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo." /clone="596372" /clone_lib="Soares mouse 3NbMS" /sex="male" /tissue_type="Spleen" /dev_stage="4 weeks" /lab_host="DH10B" <1..>523		
FEATURES	source		
	126 a	142 c	129 g 126 t
BASE COUNT	126 a	142 c	129 g 126 t
ORIGIN			
	mrna		

Db
228 gctgtgcaaggacatcgccctacaadcadagcaccatcatgcccaaccttcttggtggcacacga 287
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
COY 399 GCTGTGCCAAGGCGATCGAATACCAAGAATTCGGCTGCCCAACCTGCTGGGTCACGAGAC 458
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
288 ccagaggagcggggcttgaggtagaatcaattctaccgcgtggttaagggtgcagtgtctc 347
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
CATGAAGGAGTGCTGTGAGCAGGCAGGCGCCCTTGGATCCCCTGTGTATGAAGCAGTGTCCA 518

Db	267	gctgtgcacggacatgcgcctacaaccagaccatcatgccaaaccttctgggccacacgaa	326
QY	399	GCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCCAACCTGCTGGCCACGAGAC	458
Db	327	ccaggaggacgagccttagaggtgcaccagtctctatccgctggtgaaggtgcagtgcctc	386
QY	459	CATGAAGGAGGTCCTCGAGCAGGCCGGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCA	518
Db	387	gcccgaaactgcgtctctcctgtgtcctcatgtacgcacccgtgtgcaccg-tg--ctgga	443
QY	519	CCCGGACACCAAGAGTTCCTGTGTCTCGCTCTTTCGCCCCCGTCTGCCTCGATGACCTAGA	578
Db	444	acaggccatcccccgctgccgctctatctgtgagcgcgcgcgcccagggtcgcgaagccct	503

Query Match 2.0%: Score 41: DB 80: Length 1923;



(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

Mpsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Tue Dec 9 10:16:10 1997; MasPar time 439.22 Seconds
Tabular output not generated. 964.187 Million cell updates/sec

Title: >848439-3-trans
Description: (1-825) from translate.seq
Perfect Score: 688
N.A. Sequence: 1 WSNCGNMGNGGNYTNTTYT.....SNATNMGNAARYTNCARTGY 825
Comp: SWNCGNKCCNRRANAARRA.....WNTANKCNTYRANGTYACR

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 707517 seqs, 256659390 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: EST-STS
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
99:EST99

Database: EST-STS-TWO
100:EST100 101:EST101 102:EST102 103:EST103 104:EST104
105:EST105 106:EST106 107:EST107 108:EST108 109:EST109
110:EST110 111:EST111 112:EST112 113:EST113 114:EST114
115:EST115 116:EST116 117:EST117 118:EST118 119:EST119
120:EST120 121:EST121 122:EST122 123:EST123 124:EST124
125:EST125 126:EST126 127:EST127 128:EST128 129:EST129
130:EST130 131:EST131 132:EST132 133:EST133 134:EST134
135:EST135 136:EST136 137:EST137 138:EST138 139:EST139
140:EST140 141:EST141 142:EST142 143:EST143 144:EST144
145:EST145 146:EST146 147:EST147 148:EST148 149:EST149
150:EST150 151:EST151 152:EST152 153:EST153 154:EST154
155:EST155 156:EST156 157:EST157 158:EST158 159:EST159
160:EST160 161:EST161 162:EST162 163:EST163 164:EST164
165:EST165 166:EST166 167:EST167 168:EST168 169:EST169
170:EST170 171:EST171 172:EST172 173:EST173 174:EST174

Statistics: Mean 14.817; Variance 4.177; scale 3.547

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	345	50.1	462	174	W77968	zd70ell1.r1 Soares fet	0.00e+00
2	319	46.4	396	178	AA073862	mm97f07.r1 Stratagene	0.00e+00
3	304	44.2	426	72	H87071	ys74d07.r1 Homo sapie	0.00e+00
4	279	40.6	406	61	H16121	yl28c05.r1 Homo sapie	0.00e+00
5	277	40.3	414	156	AA024771	ze76g10.r1 Soares fet	0.00e+00
6	242	35.2	344	61	H14917	yl26d03.r1 Homo sapie	2.91e-283
7	210	30.5	299	148	W08345	mb41f02.r1 Soares mou	3.91e-237
8	163	23.7	231	65	H29095	ym31e03.r1 Homo sapie	2.17e-170
9	151	21.9	250	172	AA061047	mj78al2.r1 Soares mou	1.30e-153
10	63	9.2	531	12	M89402	CEL08B2 Caenorhabditi	1.52e-37
11	57	8.3	360	141	C09584	C.elegans cDNA clone	1.58e-30
12	52	7.6	360	89	CELK117B4F	C.elegans cDNA clone	7.26e-25
13	41	6.0	360	89	CELK079D1F	C.elegans cDNA clone	3.40e-13
14	40	5.8	256	47	H45312	yo55b03.r1 Homo sapie	3.35e-12
15	37	5.4	206	101	N63268	yz88d01.s1 Homo sapie	2.64e-09
16	37	5.4	418	101	N62714	yz76d05.s1 Homo sapie	2.64e-09
17	36	5.2	341	77	CELK119F4R	C.elegans cDNA clone	2.27e-08
18	34	4.9	336	123	HSBA0B082	H. sapiens partial cd	1.49e-06
19	33	4.8	227	184	AA017479	ze38h10.r1 Soares ret	1.14e-05
20	33	4.8	350	187	AA144540	mr17f12.r1 Soares mou	1.14e-05
21	33	4.8	414	194	AA169546	zo89b08.r1 Stratagene	1.14e-05
22	33	4.8	428	95	N41923	yy07a02.r1 Homo sapie	1.14e-05
23	33	4.8	433	194	AA169521	zo89a08.r1 Stratagene	1.14e-05
24	33	4.8	443	113	W03474	za08a10.r1 Soares mel	1.14e-05
25	33	4.8	458	23	R42865	yg06a10.s1 Homo sapie	1.14e-05
26	32	4.7	219	129	HSC0EA112	H. sapiens partial cd	8.27e-05
27	32	4.7	226	156	AA027879	zk05e04.r1 Soares pre	8.27e-05
28	32	4.7	242	119	W44804	zb98g09.s1 Soares par	8.27e-05
29	32	4.7	253	135	R58334	G2318 Fetal heart Hom	8.27e-05
30	32	4.7	262	81	T35329	EST82912 Homo sapiens	8.27e-05
31	32	4.7	384	95	N42176	yy07e11.r1 Homo sapie	8.27e-05
32	32	4.7	453	95	N42436	yy09h05.r1 Homo sapie	8.27e-05
33	32	4.7	473	127	W90442	zh78c02.s1 Soares fet	8.27e-05
34	32	4.7	584	4	T67709	yc27c11.s1 Homo sapie	8.27e-05
35	31	4.5	252	109	HUM069H03B	Human fetal brain cdn	5.73e-04
36	31	4.5	252	48	HUM055B07B	Human fetal brain cdn	5.73e-04
37	31	4.5	252	109	HUM069D03B	Human fetal brain cdn	5.73e-04
38	31	4.5	330	13	T01376	wEST02097 Caenorhabdi	5.73e-04
39	31	4.5	364	42	R27662	yh64c07.r1 Homo sapie	5.73e-04
40	31	4.5	390	1	CELK025G2R	C.elegans cDNA clone	5.73e-04
41	31	4.5	404	13	T01720	WEST02441 Caenorhabdi	5.73e-04
42	31	4.5	429	8	T85962	yd62b09.r1 Homo sapie	5.73e-04
43	31	4.5	443	144	N94473	zb80a06.s1 Soares sen	5.73e-04
44	31	4.5	574	135	N98880	zb87h03.s1 Homo sapie	5.73e-04
45	30	4.4	344	14	T09517	0113m3 Plasmodium fal	3.77e-03

ALIGNMENTS

RESULT 1 W77968 462 bp mRNA EST 17-OCT-1996
LOCUS zd70ell1.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
DEFINITION 346028 5'.
ACCESSION W77968
NID g1388502
KEYWORDS EST.

QY 516 YGAYTTYGCNYTNAARATNAARGTNAARGARATNACNTAYATNAAYMNGAYAC 569

RESULT 8 H29095 231 bp mRNA EST 17-JUL-1995

LOCUS ym31e03.r1 Homo sapiens CDNA clone 49614 5'.

DEFINITION H29095

ACCESSION 9900005

NID

KEYWORDS EST.

SOURCE human clone=49614 library=Soares infant brain lNIB vector=lafmid BA host=DH10B (ampicillin resistant) primer=M13RP1 Rsite1-Not I Rsite2-Hind III Whole brain from a 73 days post natal female. 1st strand CDNA was primed with a Not I - oligo(dT) primer [5' AACTGGAAGAATTCGGCGCGCAGGAATTTTTTTTTTTT 3']; double-stranded CDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lafmid BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo.

ORGANISM Homo sapiens

Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 231)

AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.

TITLE The WashU-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT GDB: G00-422-155

Contact: Wilson RK

WashU-Merck EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

High quality sequence stops: 127

Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES

source Location/Qualifiers

1..231

/organism="Homo sapiens"

/clone="49614"

BASE COUNT 70 a 54 c 58 g 39 t 10 others

ORIGIN

Query Match 23.7%; Score 163; DB 65; Length 231;

Best Local Similarity 55.2%; Pred. No. 2.17e-170;

Matches 128; Conservative 51; Mismatches 51; Indels 2; Gaps 2;

Db 1 cactgaaataaaantnaaggagataaacctacatcacccgagataccaaanncatcctgg 60

| : | | | | | : | | | | | | | | | | | : | | | | | : | | | |

QY 524 CNYTNAARATNAARGTNAARGARATNACNTAYATNAAYMNGAYACNAARATNATNYTG 583

Db 61 gagaccaagagcaagaccatttacaagctgaacggtgtgtccganag-gacctgaagaat 119

| : | | | | | : | | | | | | | | | | | : | | | | | : | | | |

QY 584 -ARACNAARWSNAARACNATNAYAARYTNAAYGGNGTNGARMGNGAYYTNAARAAR 642

Db 120 tcggtccctgtggctcanagacagcttgcagtgcaacctgtgnaggagatgaacgacatcaac 179

: | : | | | | : | | | | | : | | | | | | : | | | | | | | |

QY 643 WSNGTNYTNGGYTNAARGAYWSNYTNCARTGYACNTGYGARGARATGAAYGAYATNAAY 702

Db 180 gcgcccctatctggncaatgggacagacacacagggtngggagctggtgntcacct 231

| | | | | : | | | | | | | | | | | | | | | | | | | |

QY 703 GCNCCNTAYYTNGTNGTATGGGNCARAARCARCGGNGGNGARYTNGTNGTATNACNW 754

RESULT 9

LOCUS AA061047 250 bp mRNA EST 23-SEP-1996

DEFINITION mj78al2.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 482206 5'.

ACCESSION AA061047

NID g1554620

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 250)

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.

TITLE The WashU-HHMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:292950

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 179.

Location/Qualifiers

1..250

/organism="Mus musculus"

/note="Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGCGCCGCAATTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."

/clone="482206"

/clone_lib="Soares mouse p3NMF19.5"

/dev_stage="19.5 dpc total fetus"

/lab_host="DH10B (ampicillin resistant)"

<1..>250

mRNA

BASE COUNT 55 a 71 c 75 g 49 t

ORIGIN

Query Match 21.9%; Score 151; DB 172; Length 250;

Best Local Similarity 56.1%; Pred. No. 1.30e-153;

Matches 105; Conservative 48; Mismatches 34; Indels 0; Gaps 0;

Db 1 gaaatccgtgctgtggctcaaagacagcctgcagtcgcgcctgtgaggagatgaacgacat 60

: | | : : | | : | | | | | | | : : | | | | : | | | | | | |

QY 639 RAARWSNGTNYTNGGYTNAARGAYWSNYTNCARTGYACNTGYGARGARATGAAYGAYAT 698

Db 61 caacgtccgtatctgtgctcatgggacagaagcaggcgcgctggtgatgacctcctg 120

| | : | | | | | | | | | | | | | | | | | | | | | | |

QY 699 NAAYGNCNTAYYTNGTNGTATGGGNCARAARCARCGGNGGNGARYTNGTNGTATNACNWSNGT 758

: | : | | | | : | | : : | | | | | | | | : | | | | | | |

Db 121 gaaacggtggcagaaggccagagatagttcaagcgcatctcccgcagcatccgcaagct 180

| | : : | | | | | | | | : | | | | | | | : : | | : | | | |

QY 759 NAARMGNTGGCARAARGGNCARMGNGARTTYAARMGNATNWSNMGNSNATNMGNARYT 818

Db 181 gcaatgc 187

| | : | | :

QY 819 NCARTGY 825

RESULT 10

LOCUS M89402 531 bp mRNA EST 02-DEC-1992

DEFINITION CEL08B2 Caenorhabditis elegans cDNA clone cm08b2 5' similar to gene

ACCESSION M89402

NID 9275297

KEYWORDS EST.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Waterston R.H.(USA) and Sulston J.E.(UK)

(USA) Dept. of Genetics or (UK)

(USA) Washington Univ. School of Medicine or (UK) MRC Laboratory of

Molecular Biology

Box 8232,4566 Scott Ave.,St. Louis,MI 63110,USA,or,Hills

Road,Cambridge CB2 2QH,UK

Tel: (USA) (314)3627072 or (UK) (0223)248011

Fax: (USA) (314)3624137 or (UK) (0223)402008

Email: rw@ematode.wustl.edu or jes@mrc-lmb.cam.ac.uk.

Location/Qualifiers

1..531

/organism="Caenorhabditis elegans"

/clone="cm08b2"

/strain="Bristol N2"

BASE COUNT 130 a 123 c 136 g 138 t 4 others

ORIGIN

Query Match 9.2%; Score 63; DB 12; Length 531;

Best Local Similarity 37.5%; Pred. No. 1.52e-37;

Matches 118; Conservative 46; Mismatches 151; Indels 0; Gaps 0;

Db 117 atcatgagatgtgcaacgatttgccgtataacttaacgagcttcccaaatctcgtcgacg 176

||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|

QY 77 AYTNCARYTNTGTCAYGGNATNGARTAYCARAAATGMGNYTNCNAAYYTNTNGGNC 136

Db 177 aggaatcatggaagacgcctccgaatccatccatcaagccccctgctctccgttg 236

||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|

QY 137 AYGARACNATGAARGARGTNYTNGARCARGCNGGNGCNTGGATNCCNYTNGTNGTGAARC 196

Db 237 tctgctccgagcagctcaaatcttcctgtgctcctgctacttcccgatgtgcaacgaga 296

||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|

QY 197 ARTGYCAYCCNGAYACNAARAARTTYTNTGYWSNTYNTTYGCNCCNGTNGTYTNGAYG 256

Db 297 aactagccaacccaattgggtccatgccgtccattgtgcttccgtccaggaaaagtgtc 356

||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|

QY 257 AYYTNGAYGARACNATNCARCCNTGYCAYWSNTYNTGYTNCARGTNAARGAYMGNTGYG 316

Db 357 ttccagtgctggaagtttcggtttcaagtgcccgatgtgattcgtttgtgataagttcc 416

||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|

QY 317 CNCNGTNGTWSNGNTTYGGNTTYCCNTTGGCCNGAYATGYTNGARTGYGAYMGNTTYC 376

Db 417 cgtngagaacaatc 431

QY 377 CNCARGAAYAAYGAY 391

RESULT 11

LOCUS C09584 360 bp mRNA EST 09-SEP-1996

DEFINITION C.elegans cDNA clone yk173b2 : 5' end, single read.

ACCESSION C09584

NID g1534655

KEYWORDS EST(expressed sequence tag).

SOURCE Caenorhabditis elegans (strain:N2) embryo hermaphrodite cDNA to mRNA, clone lib:Yuji Kohara unpublished cDNA clone:yk173b2.

ORGANISM

Caenorhabditis elegans

Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;

Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;

Rhabditidae; Caenorhabditis.

1 (bases 1 to 360)

Kohara,Y.

AUTHORS Direct Submission

TITLE Submitted (29-APR-1996) to the DDBJ/EMBL/GenBank databases. Yuji

JOURNAL Kohara, National Institute of Genetics, Gene Library Lab; Yata

1111, Mishima, Shizuoka 411, Japan (E-mail:ykohara@lab.nig.ac.jp,

Tel:0559-81-6854, Fax:0559-81-6855)

2 (sites)

REFERENCE

AUTHORS Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A.,

Sano,M., Miyata,A. and Nishigaki,A.

TITLE Expression map of the C.elegans genome

JOURNAL Unpublished (1996)

FEATURES

Location/Qualifiers

1..360

/organism="Caenorhabditis elegans"

/strain="N2"

/clone="yk173b2"

/clone_lib="Yuji Kohara unpublished cDNA"

/dev_stage="embryo"

/sex="hermaphrodite"

BASE COUNT 86 a 90 c 90 g 87 t 7 others

ORIGIN

Query Match 8.3%; Score 57; DB 141; Length 360;

Best Local Similarity 41.9%; Pred. No. 1.58e-30;

Matches 72; Conservative 29; Mismatches 71; Indels 0; Gaps 0;

Db 181 ttctgtgctccgtctacttcccgatgtgcaacgagaaactagccaacccaattggtcca 240

||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|

QY 220 TTYTNTGYWSNTYNTTYGCNCCNGTNGTYTNGAYGAYTNGAYGARACNATNCARCCN 279

Db 241 tgccgtccattgtgtcttccgtccaggaaaagtgtcttccagtgctggaaagtctcggt 300

||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|

QY 280 TGYCAYWSNTYNTGYTNCARGTNAARGAYMGNTGYGCNCCNGTNGTWSNGCNTTYGNN 339

Db 301 ttcaagtgcccgatgtgattcgtgtgtgataaagttcccggttgaggaaacaatc 352

||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|

QY 340 TTYCCNTGGCCNGAYATGYTNGARTGYGAYMGNTTYCCNCARGAAYAAYGAY 391

RESULT 12

LOCUS CELK117B4F 360 bp mRNA EST 13-DEC-1995

DEFINITION C.elegans cDNA clone yk117b4 : 5' end, single read.

ACCESSION D76340

NID g1122125

KEYWORDS EST(expressed sequence tag).

SOURCE Caenorhabditis elegans (strain N2,) (library: Yuji kohara unpublished cDNA) Hermaphrodite embryo cDNA to mRNA.

ORGANISM

Caenorhabditis elegans

Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;

Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;

Rhabditidae; Caenorhabditis.

1 (bases 1 to 360)

Kohara,Y., Motohashi,T., Tabara,H., Sugimoto,A., Watanabe,H. and

Nishigaki,A.

AUTHORS Toward an expression map of the C.elegans genome

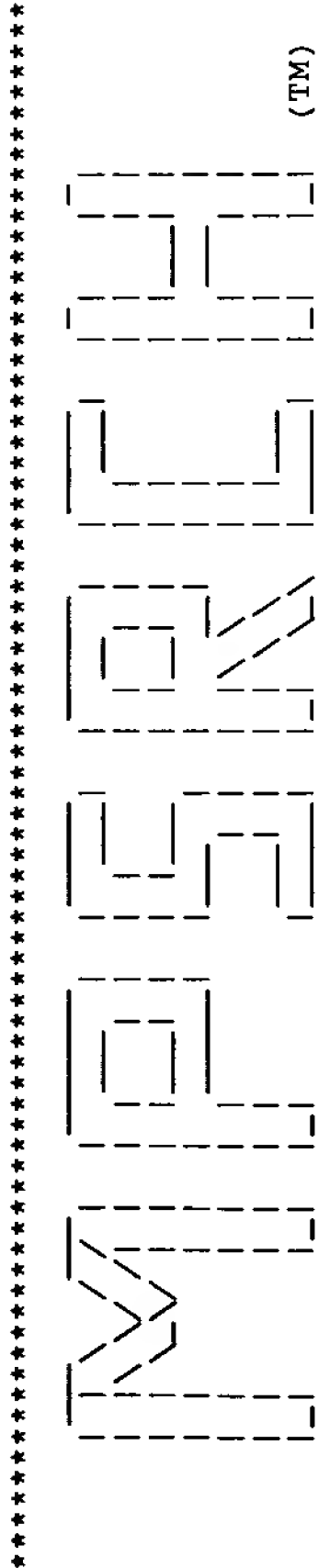
TITLE Unpublished (1995)

JOURNAL Submitted (23-Aug-1995) to DDBJ by:

COMMENT Yuji Kohara

Gene Library Lab.

National Institute of Genetics



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Tue Dec 9 10:39:19 1997; MasPar time 224.95 Seconds
1007.870 Million cell updates/sec
Tabular output not generated.

Title: >848439-3-trans
Description: (1-825) from translate.seq
Perfect Score: 688
N.A. Sequence: 1 WSNCGNMGNGGNYTNTTYT.....SNATMGNAARYTNCARTGY 825
Comp: SWNCGNKCNCNRRANAARRA.....WNTANKNTTYRANGTYACR

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0
Searched: 359085 seqs, 137405154 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: EST-STS-THREE
1:EST199 2:EST200 3:EST201 4:EST202 5:EST203 6:EST204
7:EST205 8:EST206 9:EST207 10:EST208 11:EST209 12:EST210
13:EST211 14:EST212 15:EST213 16:EST214 17:EST215
18:EST216 19:EST217 20:EST218 21:EST219 22:EST220
23:EST221 24:EST222 25:EST223 26:EST224 27:EST225
28:EST226 29:EST227 30:EST228 31:EST229 32:EST230
33:EST231 34:EST232 35:EST233 36:STS1 37:STS2 38:STS3
39:STS4 40:STS5 41:STS6 42:STS7 43:STS8 44:STS9 45:STS10
46:STS11 47:STS12 48:STS13

Database: EST-STS-FOUR
49:gnEST1 50:gnEST2 51:gnEST3 52:gnEST4 53:gnEST5
54:gnEST6 55:gnEST7 56:gnEST8 57:gnEST9 58:gnEST10
59:gnEST11 60:gnEST12 61:gnEST13 62:gnEST14 63:gnEST15
64:gnEST16 65:gnEST17 66:gnEST18 67:gnEST19 68:gnEST20
69:gnEST21 70:gnEST22 71:gnEST23 72:gnEST24 73:gnEST25
74:gnSTS 75:enEST1 76:enEST2 77:enEST3 78:enEST4
79:enEST5 80:enEST6 81:enEST7 82:enEST8 83:enEST9
84:enEST10 85:enEST11 86:enEST12 87:enEST13 88:enEST14
89:enEST15 90:enEST16 91:enEST17 92:enEST18 93:enEST19
94:enEST20 95:enEST21 96:enEST22 97:enEST23 98:enEST24
99:enEST25 100:enEST26 101:enEST27 102:enEST28
103:enEST29 104:enEST30 105:enEST31 106:enSTS 107:ueEST1
108:ueEST2

Statistics: Mean 14.890; Variance 4.327; scale 3.441
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
Result Query
SUMMARIES

No.	Score	Match	Length	DB	ID	Description	Pred. No.
1	357	51.9	458	33	AA105749	ml84a03.r1 Stratagene	0.00e+00
2	319	46.4	396	54	AA073862	mm97f07.r1 Stratagene	0.00e+00
3	235	34.2	394	101	MMAA60087	va52g08.r1 Soares mou	7.48e-265
4	210	30.5	299	94	MM3451	mb41f02.r1 Soares mou	6.10e-230
5	148	21.5	184	20	AA209833	mo90h01.r1 Beddington	6.64e-145
6	145	21.1	523	56	AA122822	mr03c11.r1 Soares mou	7.24e-141
7	145	21.1	523	6	AA122822	mr03c11.r1 Soares mou	7.24e-141
8	136	19.8	287	32	AA082155	ze88f06.r1 Soares fet	8.23e-129
9	55	8.0	502	21	AA020088	mh49d08.r1 Soares mou	1.99e-27
10	38	5.5	213	28	AA023595	mh80d03.r1 Soares mou	5.16e-10
11	37	5.4	421	82	HS1147060	zrl8a11.s1 Stratagene	4.29e-09
12	37	5.4	421	63	AA227473	zrl8a11.s1 Stratagene	4.29e-09
13	36	5.2	512	29	AA044908	zf51e10.r1 Soares ret	3.44e-08
14	35	5.1	432	29	AA030216	mh87g08.r1 Soares mou	2.66e-07
15	35	5.1	519	39	G08492	human STS CHLC.ATA11D	2.66e-07
16	34	4.9	288	36	DM57F11T	D. melanogaster STS d	1.97e-06
17	34	4.9	496	10	AA013810	mh06e01.r1 Soares mou	1.97e-06
18	33	4.8	241	21	AA019228	ze59c03.r1 Soares ret	1.41e-05
19	33	4.8	252	37	G03067	human STS WI-5108.	1.41e-05
20	33	4.8	350	56	AA144540	mr17f12.r1 Soares mou	1.41e-05
21	31	4.5	407	89	HSA62672	zs23f08.r1 Soares NbH	6.23e-04
22	31	4.5	451	39	G09558	human STS CHLC.GCT8C0	6.23e-04
23	30	4.4	177	36	DM52H6T	D. melanogaster STS d	3.85e-03
24	30	4.4	199	10	AA017992	mh45a07.r1 Soares mou	3.85e-03
25	30	4.4	213	10	AA018016	mh45c07.r1 Soares mou	3.85e-03
26	30	4.4	305	21	AA019656	ze61f11.s1 Soares ret	3.85e-03
27	30	4.4	339	43	G26024	human STS EST186690.	3.85e-03
28	30	4.4	402	1	N97824	1353C3 czappFDd2.1, D	3.85e-03
29	30	4.4	407	21	AA022630	ze71h06.s1 Soares fet	3.85e-03
30	30	4.4	425	43	G24362	human STS WI-13950.	3.85e-03
31	30	4.4	435	34	AA141629	CK02453.3prime Drosop	3.85e-03
32	30	4.4	440	26	N64746	yz30a03.s1 Soares mul	3.85e-03
33	30	4.4	455	80	CPAA24647	CpEST.119 psKIIminusC	3.85e-03
34	30	4.4	455	62	AA224647	CpEST.119 psKIIminusC	3.85e-03
35	30	4.4	471	44	G27881	human STS SHGC-34074.	3.85e-03
36	30	4.4	503	92	MM1157919	mv26g08.r1 GuayWoodfo	3.85e-03
37	30	4.4	503	70	AA241149	mv26g08.r1 GuayWoodfo	3.85e-03
38	29	4.2	170	36	DM155E2T	D. melanogaster STS d	2.26e-02
39	29	4.2	298	1	N97343	0109M3 gmbPFHB3.1, G.	2.26e-02
40	29	4.2	367	75	AT15723	22538 Lambda-PRL2 Ara	2.26e-02
41	29	4.2	386	56	AA163930	mr19c02.r1 Soares mou	2.26e-02
42	29	4.2	427	39	G09767	human STS CHLC.ATA15C	2.26e-02
43	29	4.2	449	66	AA237923	mx14f07.r1 Soares mou	2.26e-02
44	29	4.2	491	101	MMAA63807	mr18b06.r1 Soares mou	2.26e-02
45	29	4.2	522	73	C22763	Dictyostelium discoid	2.26e-02

ALIGNMENTS

RESULT	1	AA105749	458 bp	mrna	EST	04-FEB-1997
LOCUS		ml84a03.r1 Stratagene mouse kidney (#937315)				
DEFINITION		clone 518668 5' similar to TR:G1151260				
		RECEPTOR. ;				
ACCESSION		AA105749				
NID		g1654838				
KEYWORDS		EST.				
SOURCE		house mouse.				
ORGANISM		Mus musculus				
		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;				
		Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;				
		Mus.				
REFERENCE		1 (bases 1 to 458)				
AUTHORS		Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,				
		Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,				
		Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,				
		Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and				
		Waterston,R.				
TITLE		The WashU-HHMI Mouse EST Project				
JOURNAL		Unpublished (1996)				
COMMENT						

QY	336	YGGNTTYCCNTGGCCNGAYATGYTNGARTGYGAYMGNTTYCCNCARGAYAAYGAYYNTTG	395
Db	126	catcccccctcgctagtcgaccacctcctgcgcccacagagagaagctcccaaggtgtg	185
QY	396	YATNCCNYTNGCNWSNWSNGAYCAYTYNTYTCCNGCNACNGARGARGCNCNAARGTNTG	455
Db	186	tgaagcctgcataaaaccaaagaatgaggacgacacaacgacatcatcgaaaccccttgtataaaa	245
QY	456	YGARGCNTGYAARAAYAARAAYGAYGAYAAAYGAYATNATGGARACNYTNTGYAARAA	515
Db	246	tgacttcgcactgaaaaatcaaaagtgaggagagataaacgtacatcacagagacac	299
QY	516	YGAYTYGCNYTNAARATNAARGTNAARGARATNACNTAYATNAAYMNGNGAYAC	569
RESULT	5		
LOCUS	AA209833	184 bp	mrna
DEFINITION	mo90h01.r1 Beddington mouse embryonic region	Mus musculus	cdna
ACCESSION	AA209833		
NID	g1808152		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	1	(bases 1 to 184)	
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.		
TITLE	The WashU-HHMI Mouse EST Project		
JOURNAL	Unpublished (1996)		
COMMENT	Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project Washington University School of Medicinep 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:341673 Seq primer: -28ml3 rev1 ET from Amersham High quality sequence stop: 147. Location/Qualifiers 1..184 /organism="Mus musculus" /strain="C57BL6 x DBA" /note="Vector: pCMV-SPORT; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Gastrulating embryos were collected at 7.5dpc from C57BL6 x DBA matings, excluding embryos that had developed head folds and all extraembryonic tissues. Average insert size: 1.3 kb (range: 0.5 - 3.0 kb). Referenced in Development 121, 2479-2489 (1995)" /clone="560881" /clone_lib="Beddington mouse embryonic region" /sex="pooled" /dev_stage="7.5dpc" /lab_host="DH12S" <1..>184		
FEATURES	source		
BASE COUNT	61 a	45 c	48 g 30 t
ORIGIN	mrna		
Query Match	21.5%	Score 148;	DB 20; Length 184;
Best Local Similarity	58.9%	Pred. No. 6.64e-145;	
Matches	109; Conservative	45; Mismatches 30;	Indels 1; Gaps 1;

Db	1	gataacgtacatcaacagagacacccaagatcatcctggagacaaaagagcaagaccattta	60
QY	546	RATNACNTAYATNAAYMNGNGAYACNAARATNATYTNGARACNAARWSNAARACNATNTA	605
Db	61	caagctgaacggcgtgtccgaaag-gacctgaagaaaatccctgctggtcgtcaagacag	119
QY	606	YAARYTNAAYGGNGTNSNGARMGNGAYYTNAARAARWSNGTNTNGYTNAAARGAYS	665
Db	120	cctgcagtcacctgtgaggagatgaacgacatcaacgctccgtatctggtcatgggaca	179
QY	666	NYTNCARTGYACNTGYGARGARATGAAYGAYATNAAYGCNCNTAYTNGTNGTATGGNCA	725
Db	180	gaagc 184	
QY	726	RAARC 730	
RESULT	6		
LOCUS	AA122822	523 bp	mrna
DEFINITION	mr03c11.r1 Soares mouse 3NbMS	Mus musculus	cdna
ACCESSION	AA122822		
NID	g1681779		
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	1	(bases 1 to 523)	
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.		
TITLE	The WashU-HHMI Mouse EST Project		
JOURNAL	Unpublished (1996)		
COMMENT	Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project Washington University School of Medicinep 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:361804 Seq primer: -28M13 rev2 from Amersham High quality sequence stop: 469. Location/Qualifiers 1..523 /organism="Mus musculus" /strain="C57BL/6J" /note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCGCTGTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Bertrand Jordan. Library went through three rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo." /clone="596372" /clone_lib="Soares mouse 3NbMS" /sex="male" /tissue_type="Spleen" /dev_stage="4 weeks" /lab_host="DH10B" <1..>523		
FEATURES	source		
BASE COUNT	126 a	142 c	129 g 126 t
ORIGIN	mrna		


```
QY 572 ARATWATNYTNGARACNAARWSNAARACNATNTAYAARYTNAAYGGNGTNSNGARMGNG 631
Db 368 ctttaaagaagaatgaactttgg 390
      ::|||:::|:|:|
QY 632 AYYTNAARAARWSNGINYNTTGS 654

RESULT 13
LOCUS AA044908 512 bp mRNA EST 02-FEB-1997
DEFINITION zf51e10.r1 Soares retina N2b4HR Homo sapiens cDNA clone 380490 5'
            similar to contains L1.t2 L1 repetitive element ;.
ACCESSION AA044908
NID g1523130
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 512)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
TITLE WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 887 Std Error: 0.00
Seq primer: -28ML3 rev2 from Amersham
High quality sequence stop: 314.
location/Qualifiers
1..512
/organism="Homo sapiens"
/note="Organ: eye; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGGCTTTTGTTTTGTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). The retinas were obtained from a 55 year old
Caucasian and total cellular poly(A)+ RNA was extracted 6
hrs after their removal. The retina RNA was kindly
provided by Roderick R. McInnes M.D. Ph.D. from the
University of Toronto. Library constructed by Bento
Soares and M.Fatima Bonaldo."
/clone="380490"
/clone_lib="Soares retina N2b4HR"
/sex="male"
/tissue_type="retina"
/dev_stage="55 year old"
/lab_host="DH10B (ampicillin resistant)"
<1..>512

mRNA 202 a 77 c 89 g 134 t 10 others
BASE COUNT
ORIGIN

Query Match 5.2%; Score 36; DB 29; Length 512;
Best Local Similarity 37.9%; Pred. No. 3.44e-08;
Matches 47; Conservative 20; Mismatches 57; Indels 0; Gaps 0;

Db 338 taaaaattacctcaaaatagatcacacacataaaaggttaaaactataaaacttctagaag 397
      ||:||||| ||: ||||| ||||| ||||| ||||| ||||| |||||
QY 527 TNAARATWAARGTNAARGARATNCATNTAYATNAAYMGNGAYACNAARATNTYTNARA 586
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 398 aaaaatggaagaaatgtttagttactatgatgagtgaggcnaaaatttttagatggggg 457
```

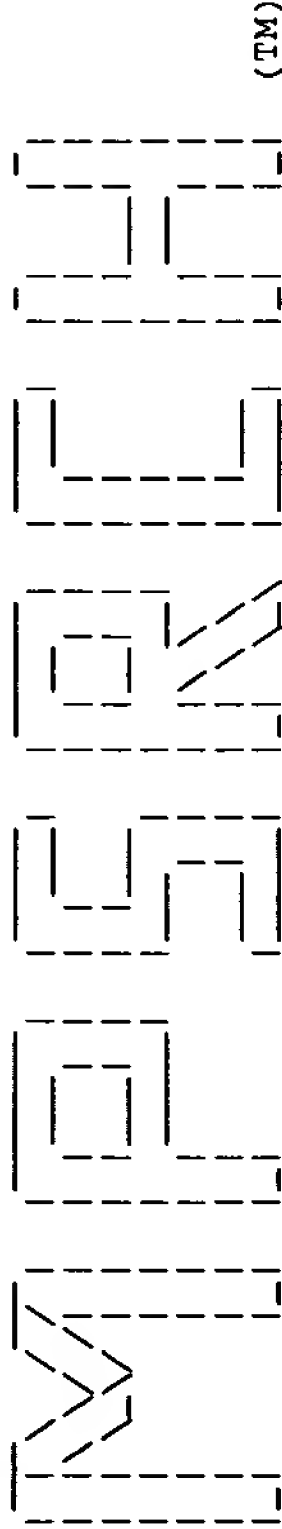
```
QY 587 CNAARWSNAARACNATNTAYAARYTNAAYGGNGTNSNGARMGNGAYYTNAARAARWSNG 646
      ||:|:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 458 tatt 461
      |:|
QY 647 TNYT 650

RESULT 14
LOCUS AA030216 432 bp mRNA EST 21-JAN-1997
DEFINITION mh87g08.r1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA
            clone 457982 5' similar to PIR:A45054 A45054 probable intercellular
            signal transducer or transmitter F2-1 ;.
ACCESSION AA030216
NID g1497354
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.
REFERENCE 1 (bases 1 to 432)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:274870
Possible reversed clone: similarity on wrong strand
Seq primer: -28MI3 rev2 from Amersham
High quality sequence stop: 399.
location/Qualifiers
1..432
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGGCGGAAATTTTGTTTTGTTTT 3'];
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M.Fatima Bonaldo."
/clone="457982"
/clone_lib="Soares mouse placenta 4NbMP13.5 14.5"
/sex="unknown"
/dev_stage="adult"
/lab_host="DH10B"
<1..>432

mRNA 75 a 137 c 153 g 67 t
BASE COUNT
ORIGIN

Query Match 5.1%; Score 35; DB 29; Length 432;
Best Local Similarity 38.8%; Pred. No. 2.66e-07;
Matches 66; Conservative 25; Mismatches 78; Indels 1; Gaps 1;

Db 50 gcacacagcgcgctacatgagcacaggaagaacttgctcgcg-gacactgcacctt 108
      :||| ||||| |||: ||||| ||||| ||||| ||||| ||||| |||||
Cp 249 RCANACNGGNGCRAANARNRNWRCANARRAAYTTYTTTNGTRTCNGGRTGRCAYTGYTCAT 190
```



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Tue Dec 9 07:57:40 1997; MasPar time 1674.10 Seconds
Tabular output not generated. 1329.789 Million cell updates/sec

Title: >US-08-848-439-1
Description: (1-2027) from US08848439.seq
Perfect Score: 2027
N.A. Sequence: 1 GAATTCGGCCTTCATGGCCT.....AAAAAAGGCGCGCGC 2027
Comp: CTTAAGCCGGAAGTACCGGA.....TTTTTTTTTTCGCGCGCG

Scoring table: TABLE default
Gap 6
Nmatch STD : Dbase 0; Query 0
Searched: 362067 seqs, 549138275 bases x 2
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb1-new3
1: BCT 2: FUN 3: GEN1 4: GEN2 5: HTG1 6: HTG2 7: HUM 8: INV
9: ORG 10: MAM 11: VRT 12: PLN 13: PRO 14: ROD 15: SYN 16: UNC
17: VTR
Database: genbank99
18: BCT1 19: BCT2 20: BCT3 21: BCT4 22: BCT5 23: BCT6 24: BCT7
25: BCT8 26: BCT9 27: BCT10 28: BCT11 29: GEN1 30: GEN2
31: GEN3 32: HTG1 33: HTG2 34: HTG3 35: INV1 36: INV2 37: INV3
38: INV4 39: INV5 40: INV6 41: INV7 42: INV8 43: INV9 44: INV10
45: INV11 46: MAM1 47: MAM2 48: MAM3 49: VRT1 50: VRT2 51: VRT3
52: VRT4 53: PAT1 54: PAT2 55: PAT3 56: PAT4 57: PAT5 58: PHG
59: PLN1 60: PLN2 61: PLN3 62: PLN4 63: PLN5 64: PLN6 65: PLN7
66: PLN8 67: PLN9 68: PLN10 69: PLN11 70: PRI1 71: PRI2
72: PRI3 73: PRI4 74: PRI5 75: PRI6 76: PRI7 77: PRI8 78: PRI9
79: PRI10 80: PRI11 81: PRI12 82: PRI13 83: PRI14 84: PRI15
85: ROD1 86: ROD2 87: ROD3 88: ROD4 89: ROD5 90: ROD6 91: ROD7
92: ROD8 93: STR 94: SYN 95: UNA 96: VRL1 97: VRL2 98: VRL3
99: VRL4 100: VRL5 101: VRL6 102: VRL7 103: VRL8 104: VRL9
105: VRL10
Database: genbank-new3
106: BCT 107: GEN1 108: GEN2 109: HTG1 110: HTG2 111: INV
112: MAM 113: VRT 114: PHG 115: PLN 116: PRI1 117: PRI2
118: ROD 119: SYN 120: UNA 121: VRL
Database: u-emb150_99
122: part1

Database: genbank-new3
106: BCT 107: GEN1 108: GEN2 109: HTG1 110: HTG2 111: INV
112: MAM 113: VRT 114: PHG 115: PLN 116: PRI1 117: PRI2
118: ROD 119: SYN 120: UNA 121: VRL
Database: u-emb150_99
122: part1

Statistics: Mean 12.293; Variance 6.329; scale 1.942
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	959	47.3	1776	85	D50462	Mouse SDF5 mRNA, comp	0.00e+00
2	58	2.9	2334	77	HSU43318	Human putative transm	5.31e-19
3	56	2.8	2421	87	MMU43321	Mus musculus putative	1.17e-17
4	45	2.2	1912	90	RATFRZH	Rattus norvegicus Dro	1.59e-10
5	41	2.0	1923	80	HUMFRIZ	Human frizzled gene p	4.63e-08
6	39	1.9	215	57	I28278	Sequence 5 from paten	7.33e-07
7	38	1.9	4540	90	RATFRZ2H	Rattus norvegicus Dro	2.85e-06
8	31	1.5	215	57	I28278	Sequence 5 from paten	2.40e-02
9	30	1.5	354	111	OFU89259	Oxytricha fallax 57kD	8.07e-02
10	30	1.5	354	8	OFU89259	Oxytricha fallax 57kD	8.07e-02
11	30	1.5	4297	59	ATHSC701	A.thaliana hsc70-1 ge	8.07e-02
12	31	1.5	38262	37	CELK02A2	Caenorhabditis elegan	2.40e-02
13	28	1.4	354	8	OFU89259	Oxytricha fallax 57kD	8.46e-01
14	28	1.4	354	111	OFU89259	Oxytricha fallax 57kD	8.46e-01
15	28	1.4	565	73	HSANX5S11	Human annexin V (ANX5	8.46e-01
16	28	1.4	616	73	HSANX5V11	Human annexin V (ANX5	8.46e-01
17	29	1.4	1073	59	ATCOR47	A.thaliana cor47 mRNA	2.65e-01
18	28	1.4	1444	79	HUMATC	Human placental antic	8.46e-01
19	28	1.4	1454	55	I08832	Sequence 3 from Paten	8.46e-01
20	28	1.4	1466	53	A07367	M.musculus VAC-alpha	8.46e-01
21	28	1.4	1560	78	HSVAC	Human mRNA for vascul	8.46e-01
22	28	1.4	1566	79	HUMBCI	Human blood coagulati	8.46e-01
23	28	1.4	1567	57	I33410	Sequence 2 from paten	8.46e-01
24	28	1.4	1573	44	RFAT2A711	R.filosa mRNA for alp	8.46e-01
25	28	1.4	1575	53	A13845	H.sapiens PP4 mRNA.	8.46e-01
26	28	1.4	1582	82	HUMPAP4	Human placenta antico	8.46e-01
27	28	1.4	1592	80	HUMENN	Human endonexin II MR	8.46e-01
28	28	1.4	1605	55	I07181	Sequence 12 from Pate	8.46e-01
29	29	1.4	2028	78	HSU56813	Human polycystwin MRN	2.65e-01
30	29	1.4	2112	85	MMEBK	M.musculus mRNA for E	2.65e-01
31	29	1.4	2260	87	MMU43320	Mus musculus putative	2.65e-01
32	28	1.4	2290	43	PFAHSP	Plasmodium falciparum	8.46e-01
33	28	1.4	3205	22	ECOUVRAA	E.coli uvrA gene enco	8.46e-01
34	29	1.4	8595	39	DDU14576	Dictyostelium discoid	2.65e-01
35	28	1.4	10419	21	ECAE000479	Escherichia coli from	8.46e-01
36	28	1.4	176195	22	ECOUW89	E. coli chromosomal r	8.46e-01
37	27	1.3	105	56	I14734	Sequence 13 from pate	2.63e+00
38	27	1.3	1539	61	CRCPRFROM	C.roseus cpr gene, pr	2.63e+00
39	27	1.3	1738	60	ATU53866	Arabidopsis thaliana	2.63e+00
40	27	1.3	2147	63	NTPKTL7	N.tabacum PKTL7 mRNA	2.63e+00
41	27	1.3	3215	35	AGRNAG1	A.gambiae mRNA for im	2.63e+00
42	27	1.3	4203	85	MM5HT3RCB	M.musculus 5HT3 serot	2.63e+00
43	27	1.3	8631	44	PFPOLII	Plasmodium falciparum	2.63e+00
44	27	1.3	27694	62	HASMT	Hansenula wingei mito	2.63e+00
45	27	1.3	62640	32	CEF26H11	*** SEQUENCING IN PRO	2.63e+00

26697

log(a)or(b)

ALIGNMENTS

RESULT 1
LOCUS D50462 1776 bp mRNA ROD 19-DEC-1996
DEFINITION Mouse SDF5 mRNA, complete cds.
ACCESSION D50462
NID gl747301
KEYWORDS SDF5.
SOURCE Mus musculus cell_line:ST-2 cdna to mRNA, clone_lib:phage (lgt22a)
library, ST-2 directional 1.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.
REFERENCE 1 (bases 1 to 1776)
AUTHORS Shirozu,M.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-1995) to the DDBJ/EMBL/GenBank databases. Michio Shirozu, Kyoto University, Faculty of Medicine, Department of Medical Chemistry; Yoshida, Sakyo-ku, Kyoto, Kyoto 606, Japan (E-mail:kondo@virus1.virus.kyoto-u.ac.jp, Tel:81-75-753-4377, Fax:81-75-753-4388)
REFERENCE 2 (bases 1 to 1776)

Eukaryotae; mitochondrial eukaryotes; Viridiplantae;
Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
Magnoliopsida; Capparales; Brassicaceae; Arabidopsis.
1 (bases 1 to 4297)
King,K.
Unpublished
JOURNAL
AUTHORS
2 (bases 1 to 4297)
King,K.
Direct Submission
TITLE
JOURNAL
Submitted (22-JAN-1994) K. King, Dept of Genetics, Universitaet
Tuebingen, Auf der Morgenstelle 28, 72076 Tuebingen, FRG
FEATURES
source
/organism="Arabidopsis thaliana"
/strain="Ostwestfalen"
/tissue_type="leaf"
/clone="Athsc70g"
<1..171
/gene="Hsc70-1"
/number=1
join(<1..171,496..2238)
/gene="Hsc70-1"
/codon_start=1
/product="heat shock cognate 70-1"
/db_xref="PID:g450880"
/db_xref="SWISS-PROT:P22953"
/translation="GTTYSCVGVWQHDRVVEIIANDQGNRTTPSYVAFDTSERLIGDAA
KNQVAMSPVNTVFDAKRLIGRRFSDSSVQSDMKLWPFKIQAGPADKPMIYVEYKGEK
EFAAEIISVMLIKREIAEAYLGVTIKNAVTVTPAYFENDSQROATKDAGVIAGLNVN
RIINEPTAAAIAYGLDKKATRVGEKNVLIFDLGGTFDVSLLTIEEGIFEVKATAGDT
HLGGEDFDNRNVNHFVQEFKRSKKDITGNPRALRLRTSCERAKRTLSSSTAQTTEI
DSLYEGIDFYSTITRAREEELNMDLFRKCMPEVKLURDAKMDKSTVHDVVLVGGSTR
IPKVQQLQDFFNGKELCKSINPVEAVVYGAAVQAILSGEGNEKVQDLLLLLVTPLS
LGLETAGGVMTLIPRNTTIPTKKEQVFSTYSDNQPGVLIQVYEGERARTKDNLLGK
FELSGIPAPRGVQITVCFDIDANGILNVSAEDKTGQKNKVTITNDKGRLSKDEIE
KMVQAEKYKSEDEHKKVEAKNALENAYNNRNTIQDEKIGELKPAADKKKIEDSI
EQAIQWLEGNQLAEADFEEDKMKLESI CNPIIAKMYQGAGGAGGPGASGMDDDAPP
ASGGAGPKIEVD"
172..495
/gene="Hsc70-1"
/number=1
496..>2238
/gene="Hsc70-1"
/number=2
<3695..3908
/gene="Hsc70-2"
/note="see also M23106"
/number=1
join(3695..3908,4217..>4297)
/gene="Hsc70-2"
/note="see also M23106"
/codon_start=1
/product="heat shock cognate 70-2"
/db_xref="PID:g450881"
/db_xref="SWISS-PROT:P22954"
/translation="MAGKGKGAIGIDLGTTSYCVGVWQHDRVVEIIANDQGNRTTPSY
VAFDTSERLIGDAAKNQVAMNPVNTVFDAKRLIGRRFSDASVQSDRLWPFETII"
3909..4216
/gene="Hsc70-2"
/note="see also M23106"
/number=1
4217..>4297
/gene="Hsc70-2"
/note="see also M23106"
/number=2
BASE COUNT 1172 a 840 c 906 g 1379 t
ORIGIN
Query Match 1.5%; Score 30; DB 59; Length 4297;
Best Local Similarity 65.0%; Pred. No. 8.07e-02;
Matches 65; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
Db 2982 gcttatttttttttttgggtcaaccttgatgctttggttatagcttggttgat 3041

||||| ||||||| ||||||| ||| ||| | ||||| | ||||| ||| |
Cp 2021 GCTTTT TTTT TTTT TTTT TAGGTAAACAGGATGTAAAGTTTATATACAAGATATAATG 1962
Db 3042 ttatctggattttttactttgacgttttatgtttatgctttattt 3081
||||| ||| ||||| ||| ||| ||| ||| |||||
Cp 1961 TTTATCTGAAATATTACAGTGTGGTTAAAGCAATATTT 1922
RESULT 12
LOCUS CELK02A2 38262 bp DNA INV 13-JUL-1995
DEFINITION Caenorhabditis elegans cosmid K02A2.
ACCESSION U23171
NID g726381
KEYWORDS
SOURCE Caenorhabditis elegans strain=Bristol N2.
ORGANISM Caenorhabditis elegans
Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;
Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;
Rhabditidae; Caenorhabditis.
REFERENCE 1 (bases 1 to 38262)
AUTHORS Wilson,R., Ainscough,R., Anderson,K., Baynes,C., Berks,M.,
Bonfield,J., Burton,J., Connell,M., Copsey,T., Cooper,J.,
Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A.,
Fulton,L., Gardner,A., Green,P., Hawkins,T., Hillier,L., Jier,M.,
Johnston,L., Jones,M., Kershaw,J., Kirsten,J., Laister,N.,
Latreille,P., Lightning,J., Lloyd,C., McMurray,A., Mortimore,B.,
O'Callaghan,M., Parsons,J., Percy,C., Rifken,L., Roopra,A.,
Saunders,D., Shownkeen,R., Smaildon,N., Smith,A., Sonhammer,E.,
Staden,R., Sulston,J., Thierry-Mieg,J., Thomas,K., Vaudin,M.,
Vaughan,K., Waterston,R., Watson,A., Weinstein,L.,
Wilkinson-Sproat,J. and Wohldman,P.
TITLE 2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans
JOURNAL Nature 368 (6466), 32-38 (1994)
MEDLINE 94150718
REFERENCE 2 (bases 1 to 38262)
AUTHORS Hallsworth,K.
TITLE The sequence of C. elegans cosmid K02A2
JOURNAL Unpublished (1994)
REFERENCE 3 (bases 1 to 38262)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (13-JUL-1995)
COMMENT Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
e-mail: rw@nematode.wustl.edu and jesesanger.ac.uk

NEIGHBORING COSMID INFORMATION:

The 5' cosmid is C07D10, 200 bp overlap;3' cosmid is T19C12.
Actual start of this cosmid is at base position 197 of CELK02A2;
actual end is at bp 38262 of CELK02A2

NOTES:

Coding sequences below are predicted from computer analysis, using
the program Genefinder (P. Green and L. Hillier, ms in
preparation).

FEATURES
source
Location/Qualifiers
1..38262
/organism="Caenorhabditis elegans"
/map="II"
/strain="Bristol N2"
join(2467..2528,2604..2700,2748..2885,3030..3083,
3384..3656,3705..3804,4217..4404)
/gene="K02A2.2"
CDS

/note="similar to C. elegans protein E04F6.1"
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g726382"
/translation="MYGVCFHCKNLNKTEKIKNISLPNIDSPAVLPRGPCREFGPDV
CFSAYHFFLGVALCVGLISNTIIFRFQALRKGRVSRSRIFTMISLTYIPSSITIVYL
NYFLRIKNNLTMISQILPFTSSWDFEKVRSLTYYIEHPKYDLSIYEPFVGFHNIAFQV
TLATLLLVIGAYAIIPAISGFLTISRVIHLINDNRGMSLRTKEOKTLAYGLACQTFLPV
ICYIPVASCYIFSQMTSVELLNEHLLGILICFPFSDVPFISFYFIVPYRQALLGLFK
CRKTKPLNIVITVRHLSKRNPISVDSN"
complement(join(4846..4918,4971..5077))
/gene="K02A2.5"
CDS
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g726384"
/translation="MLSDDFSLSYDYDDKADAHDAADAAFRKFEALCSARSQPSTSK
PSLIDALCTSALKE"
complement(join(19741..19827,19873..20071,20408..20658,
20735..21024,21067..21346,21423..21621,21674..21793,
21924..22110,22353..22900,22950..23072,23118..23551,
23600..23743,23809..23993,24055..24070))
/gene="K02A2.3"
/note="similar to bumetanide-sensitive Na-K-Cl
cotransporter (PIR:A53491)"
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g726386"
/translation="MGKGGGRMETVGEDGIPADYKGNRKFTTSLGHLALYKEDEGIGT
QASFISGYTTPGPKERTSEHKVANKLGMVGVLPTIQIHILGVTMFIRLFVVGMSGV
ATWMAIIAICCLSTLLTSISLAVATNGVVESSGGAYFIISRLNGAEFGSAVGILFYLA
NTVAASMYIVGVEVILMYLWPEMAIGGADALHDTMEFGSLYNNLRLYGTVFLLIQAL
IVANGVKFVQLLAPVSLMCVILAAACIGGGIEKQITMEGMKVCAIDNHLLOQSSIVTH
PIHKNQTSWFNETVDFCNLCKSLYLESVFECANVNDEASAEDDVFECTHYTSKMTCTQ
LAFPGNMKTLLDNMMWPEYMEKSEVPVPGVRGKETAEVVQDESSTFFMLMAIYFPVATG
IFTGTNMSGDLRDPQRSIPVFETIAATLTTSAIYYILAILFGSITRSVLRDKFGRSIG
NTMVAAALSWPHAVVTVGAFSTFGAALQCLCSAPRLQLQSIKDDVIPILAPFARVT
KNEPFLGLVTVIIAECGILLGAVDKIAEVLDFEFLMTLSLLGAALCFFIMFASSVP
LACIACTATAVIYKYVEWKAKKEWGDGMRGLATTAQYSLKVEDKDPHPKNWRPOV
LILLTSQWSKEMIDRRAYVSMNLGAQLKAGRLAIACAFKGSVDSQDKNRRADVKT
TLVKDMSSVRLRGFAKTMFYNNHQINGTISGLYQSIGGLRPNTILLNWPNEKNPDE
LVLFABEIIHGAANDNCLIVTKGITDFPEYSERLTGFDIWIWIVQDGGILMLIAYLLR
QHKVWKGCTLRIFAVSEQDSTKSEDMKAGLQKIYMLRIDAELEFIVDLLDMEVSDEV
EKAAEVERKQKEREEMRRSKGYLNDGFMEDNGKPRQVMRHSDSARSFSPQPGAHTS
INLDETETSFTESLFDDFYRSGTPNEDLEGAMKLNHMKNTSVRLNRVIRENSPDSQL
ILLNLPSPPNRNLAFNNSYMTYLDVLTEDLPRVLFIGSGGREVITIDS"
join(26353..26567,26616..26712,26757..26894,28725..29289,
29341..29441)
/gene="K02A2.1"
/note="similar to C. elegans protein E04F6.1"
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g726385"
/translation="MSDRSNYSNRFVSYVNAYFHFFLVAGLTFQCILLYLIRTKSPAS
LDQLKYFLYNTAFVQILVIICGYFTQHRSLPNSTTFVAVLANGPCRVFGTTVCFGGYHV
FLGVTFCVALSISNTVIYRFLILRRDITKKHLIYIIILSYVPGIITTYAPLILISFN
LVSVELTDEVTFYFEKLSKLLAQEHGYVALVTDIYGKIRCTDISSAVTLLRPMTSDR
NCKLKRLEAALNALKSVPVCKQKLGAFGFCIGGLCSLDCARYRFDGIRAVISFHGT
LTPIEGIPLELLDDNSIQVHHGDADKHVSKVTVDFAFHEEMRAFNDSFVFISHGKAMHS
FTDPESAGIAPSGVGYDANAERKSWKATLEFLKEAFA"
complement(join(29971..30119,30168..30403,30450..30655))
/gene="K02A2.4"
/note="coded for by C. elegans cDNA CEESP19F; similar to
C. elegans protein E04F6.1"
/codon_start=1
/db_xref="PID:g726383"
/translation="MADDSNRFVLYFDVYYYSYFIIGVFAQSLLYLIIKNKSPRLYS
FCYFLINTWIVQFAVLLMTFFQSRCLPNSTTYTVLPRGPCYFGPTACFAGYHISLA
VSMVAVALSFANTVLFRLPIRPHGFNKNLFFDDDCNMLHSSNISCDCTFSGHLGFSKS
TYLHLPRTSDVREFFNIRTISLVFRISNQHFSLRL"
complement(join(33175..33695,33825..35082,35609..37636))
/gene="K02A2.6"
/note="similar to silkwork transposon mag (PIR:S08405) and
POL polyprotein (reverse transcriptase)"

/codon_start=1
/evidence=not_experimental
/db_xref="PID:g726387"
/translation="MNDLKRSIQKFQFNVAEPEAFKRWLARNKLTFFVEDGKNLSERER
TRLLGLCEESTFHRYESQREISDIYSISFDDTVTALTALTFEGSKSLMMRRQQCLOQI
CRANGLSQDYLDYNTSISDAVLSKSSMTSDWSIFLFRGLNSPGDEKAKLYLMQY
VEASEKNEKLLKLSDVHDEWMKFIQMHQQSKIIVSKPSKSSQQVDVNVKVDVNTNRSKKK
KPIPRPEKSSQDSKKKGEIPTCFCNKKGHYATNCRSNPKTGNQGNKGKSKGCDVS
HVDGLDVKTEHQAKHRMSVEVCCKVAFQLDTGSMITLISVKCWEKLGSPPLEKVPVHR
ISCANGTPMAVKGRCVLKFKLGEIYEVYVDRDTRNLLGTSWNLCPQMRSAQAQI
VNQVSTSETEASRLLEVMLKNDPPEVKDGLGLCTKEKAERTEENAVPVFKRARVPVY
GSLEAVETELNLQEMGVIVPITYAKWAAPIVVKKKGTKIRVCADFKCSGLNAALK
DEFHPLTSEDIFSRLLKGTVYSQIDLKDAYLQVELDEEAQKLAVINTHRGIKFLYLRMT
FGLKPAPASFQKIMDKMVSGLTGVAVYWDIIISASSIEEHEKILRLEFKEFGYGR
VSAEKCAFAQKQVTFLGFVDEHRRPDSKTEAIRSMKAPTQKQLASFGLGAADWLSR
MMQDHQONADDVIAEIIYDDDDDEDSIIQKLNPTETDIRFESQKDEHVSVVKLVR
NDSWKPKPSTEIEKHWIRYRDRRLKLIHGCLLDDRVIVPKSLQKIYKQLHEGHPGIV
QMKQKARSFVFWRGDSDIENMVHCNQCENSKMPRVPLNPWPPEAPWKRIHIDF
AGPLNGCYLLVVVDAKTKYAEVKLRSISAVTTIDLLEEIFSIGHYPETIISDNGTQL
TSHLFAQMCQSHGIEHKTSAYVYPRSNGAARFVDTLKRGIKAKIAGEGVSNNQIILNKF
LISYRNTPHSALNGSTPAECHFGKIRITMSLLMPTDRVLKVPKLTQYQQNMKHHYEL
RNGARAKAFQVNQKVVYVQVHHGNKQWKHGVIRRKFGVLYEVQVQDRMQKSHVNQIR
TRYGDIYSSSVNRLSSDPLGFVSGGDEMSRNDPVTNSGSSITDYNRGSRTISEL
KKKESNAVRGQPCGSCSPTNNDVSAPGYASRSHPPTDPSHSVRRSSIRRVLDTRYGSS
VEHPSTSTGTPRGSTSTQLGQASTRNGSRYTASGRNPSCQGNRYSSIRGEGVTARRER
VRTTWR"
BASE COUNT 12820 a 6635 c 6904 g 11903 t
ORIGIN
Query Match 1.5%; Score 31; DB 37; Length 38262;
Best Local Similarity 73.1%; Pred. No. 2.40e-02;
Matches 49; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
Db 20220 aaagcgaacttcacgaaacttatagaaataatacaataaactgagaattt 20279
||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||||
QY 1346 AAAGGAAAAGCCCCACCGAATCTTGTAGAAATATTCAAACATAATAATCATGAATATT 1405
Db 20280 ttctcga 20286
|| |||
QY 1406 TTATGAA 1412
RESULT 13
ID OFU89259 standard; DNA; INV; 354 BP.
AC U89259;
NI g1881675
DT 13-MAR-1997 (Rel. 51, Created)
DE 13-MAR-1997 (Rel. 51, Last updated, Version 1)
DE Oxytricha fallax 57kD zinc finger/protein chimera gene, partial
DE cds.
KW Oxytricha fallax
OS Eukaryotae; mitochondrial eukaryotes; Alveolata; Ciliophora;
OC hypotrichs; Stichotrichida; Oxytrichidae; Oxytricha.
RN [1]
RP 1-354
RX MEDLINE; 94134747.
RA Doak T.G., Doerder F.P., Jahn C.L., Herrick G.;
RT "A proposed superfamily of transposase genes: transposon-like
RT elements in ciliated protozoa and a common 'D35E' motif";
RL Proc. Natl. Acad. Sci. U.S.A. 91:942-946(1994).
RN [2]
RP 1-354
RA Witherspoon D.J., Doak T.G., Williams K., Seger J., Herrick G.;
RT "Selection on the protein-coding genes of the TBE1 family of
RT transposable elements in the ciliates Oxytricha fallax and O.
RT trifallax";
RL Unpublished.
RN [3]
RP 1-354
RA Doak T.G., Williams K., Witherspoon D.J., Herrick G.;
RT ; Submitted (11-FEB-1997) to the EMBL/GenBank/DBJ databases.
RL


```
/note="lipocortin V, endonexin II, PAP-I, PP4, VAC-alpha"
/codon_start=1
/product="annexin v"
/db_xref="PID:g430966"
/translation="MAQVLRGTVTDFPGFDERADAETLRKAMKGLGTDEESILTLTTS
RSNAQRQEISAAFKTLFGRDLDDDLKSELTGKFEKLIVALMKPSRLYDAYELKHALKG
AGTNEKVLTEIIASRTPEELRAIKQVYEEYGSSEDDVVGDTSYYQRMVLLQAN
RDPDAGIDEAQVEQDAQLFQAGELKWTDEKFTIFGTRSVSHLRKVFDKYMVTISG
FQIEETIDRETSGNLEQLLAVVKSIRSIPAYLAETLYAMKGAGTDDHTLIRVMVSR
SEIDLFNIRKEFRKNFATSLYSMIKGDTSGDYKKALLLCGEDD"
intron
<1..30
/evidence=experimental
/gene="ANX5"
/note="full size 0.8 kb; phase 0 intron"
/number=12
exon
31..565
/evidence=experimental
/gene="ANX5"
/number=13
3'UTR
91..565
/evidence=experimental
/gene="ANX5"
538..543
polyA_signal
/gene="ANX5"
BASE COUNT 163 a 107 c 104 g 191 t
ORIGIN
Query Match 1.4%; Score 28; DB 73; Length 565;
Best Local Similarity 80.4%; Pred. No. 8.46e-01;
Matches 37; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Db 345 atatatttttaagaggctagagtgccttttagcccttttttaaaa 390
||||| ||||| ||| | || |||| ||||| ||||| |||||
QY 1794 ATATAAATTTTATAATTGACTTGACTTAAAGCCTTGTTTAAAA 1839
```

Search completed: Tue Dec 9 08:36:17 1997
Job time : 2317 secs.

RMACEHLPLHGDPDNLMEQPSYTEAGSGSGSGSGSGSGSGSGGKRGKQKQSGSGGS
GSGSGSTSTKPCRGRNSKNQNPOGEKASGKECSCSCRSPLIFLQKEQLQOQSOM
PMHHPHWYMLNLTQRIAGVPCNGIPCKGPFNSDEKDFAGLWIALWSGLCFCSOTLM
TLTTFIIDTEREYKPERPIVFLSACYFMAVAVGLSRNFLQNEEIIACDGLLLRESSTGP
HSCTLVFLTYFFGMASSIWWVILFTFTWFLAAGLKWGNEAITKHSQYFHLAAWLIPTV
QSAVALLLSAVDGDPIIGICYVGNLNPDLKTFVLAPLFVYLVIIGTTFELMAGFVSLFR
IRSVIKQGGVGAGVKADKLEKLMIRIGIFSVLTVPATIVIGCYLYEAAYPEDWIKAI
LAPCAQOVKPGKPLYSVLMLKYFMALAVGITSVWISGKTLESWRRFRWRLLAGP
DRTGANOALIKQRPPIPHPYAGSGMGMPVGSAGSLLATPYTQAGGASVASTSHHLH
HVLKQPAASHV"

BASE COUNT 404 a 663 c 612 g 406 t
ORIGIN

Query Match 8.9%; Score 61; DB 40; Length 2085;
Best Local Similarity 38.5%; Pred. No. 7.21e-11;
Matches 120; Conservative 42; Mismatches 150; Indels 0; Gaps 0;

Db 206 taccaatgtgtcggggcattggctacaacatgacatccttccccacgaatgaaccatg 265
| | : | | | : | | | : | | : | | | : | | | : | | : |
QY 80 TNCARYTNTGYCAYGGNATGARTAYCARAAAYATGMGNYTNCNAAYYTNTNGNCAYG 139
| | : | | | : | | | : | | : | | | : | | | : | | : |
Db 266 agacccaggacgaagcggcctggagggtgcaccagtcttgccccctggaggatcaaat 325
| | : | | | : | | | : | | : | | | : | | | : | | : |
QY 140 ARACNATGAARGARTNYTNGARCAGCNGGNCNTGGATNCCNTNGTNATGAARCART 199
| | : | | | : | | | : | | : | | | : | | | : | | : |
Db 326 gctgcgcggacctaagttcttcctgtgcagcatgtacacgccatctgcctggaggatt 385
| | : | | | : | | | : | | : | | | : | | | : | | : |
QY 200 GYCAVCNGAYACNAARAARTTYTNTGYWSNYTNTTYGCNCCNGTNGYTTNGAYGAYY 259
| | : | | | : | | | : | | : | | | : | | | : | | : |
Db 386 accacaaagcgtgccccgtttgcggagtgctgcgagagagccccctcgggatgcgcac 445
| | : | | | : | | | : | | | : | | : | | | : | | | : | | |
QY 260 TNGAYGARACNATNCARCCNTGYCAYWSNYTNTGYTNCARGTNAARGAYMGNTGYGCNC 319
| | : | | | : | | | : | | : | | | : | | | : | | : |
Db 446 ccatcatgcagcagtacagcttcgaatggccggagagaatggcgtgcgagcacttgcccc 505
| | : | | | : | | | : | | | : | | | : | | : | | | : | | |
QY 320 CNGTNATGWSNGCNTTYGGNTTYCCNTGGCCNGAYATGYTNGARTGYGAYMGNTTYCCNC 379
| | : | | | : | | | : | | : | | | : | | | : | | : |
Db 506 ttcatggtgacc 517
| : | | : |
QY 380 ARGAYAAYGAYY 391

RESULT 7 MMU43320 2260 bp mRNA ROD 24-FEB-1996
LOCUS Mus musculus putative transmembrane receptor (frizzled 7) mRNA,
DEFINITION complete cds.
ACCESSION U43320
NID g1151257
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 2260)
AUTHORS Wang,Y., Macke,J.P., Abella,B.S., Andreason,K., Worley,P.,
Gilbert,D.J., Copeland,N.G., Jenkins,N.A. and Nathans,J.
TITLE A large family of putative transmembrane receptors homologous to
the product of the Drosophila tissue polarity gene frizzled
JOURNAL J. Biol. Chem. 271 (8), 4468-4476 (1996)
MEDLINE 96224032
REFERENCE 2 (bases 1 to 2260)
AUTHORS Abella,B., Wang,Y., Macke,J.P. and Nathans,J.
TITLE Direct Submission
JOURNAL Submitted (14-DEC-1995) Jeremy Nathans, Molecular Biology and
Genetics, Johns Hopkins Medical School, 725 N. Wolfe Street,
Baltimore, MD 21205
FEATURES Location/Qualifiers
source 1..2260
/organism="Mus musculus"
/chromosome="2"
/map="2q33-34"
362..2080
CDS

/gene="frizzled 7"
/note="putative transmembrane receptor"
/codon_start=1
/product="transmembrane receptor"
/db_xref="PID:g1151258"
/translation="MRGPGTAASHSPGLCALVLALLGALPTDTRAQPHYHGEKGISVP
DHGFCQPISSIPCLTDIAYNQITILPNLLGHTNOEDAGLEVHQFYPLVKVQCSPELRFEL
CSMYAPCVTDQAIPPCRSCLCERARQGCCEALMNKFGQFWPERLRCENFPVHGAGEIC
VGQNTSDGSGAGGSTAYPTAPYLPDPPTAMSPDSGRGLRSLFFSCPRQKVPYPYL
GYRFLGERDCGAPCEPRANGLMYFKEEERFRALWVGVWSVLSCASTLFTVLTYLVD
MRFSYPERPIIFLSGCYFMVAHVAGFLLEDRAVCVERFSDDGYRTVAQGTKEGC
TILFMVLYFFGMASSIWWVILSLTWFLAAGMKWGEAIEANSQYFHLAAWAVPAYVTI
TILAMGQVDGDLLSVCYVGLSSVDALRGFVLAPLFVLYFIGTSFLLAGFVSLFRIT
IMKHDGQTKTEKLEKLMVRIGVFSVLYTPATIVAPLACIFYEQAFREHWERTWLLQTKS
YAVPCPPRHFSMPSPDFTVEMIKYLMITMIVGITTGFWISGKTLQSWRRFYHRLSHSS
KGETAV"

BASE COUNT 362 a 739 c 677 g 481 t 1 others
ORIGIN

Query Match 8.6%; Score 59; DB 87; Length 2260;
Best Local Similarity 39.6%; Pred. No. 5.54e-10;
Matches 118; Conservative 44; Mismatches 133; Indels 3; Gaps 2;

Db 522 tcccggtgtgcacggatatgcctacaacacagaccatcctgcccaacctgctgggccaca 581
| | : | | | : | | | : | | : | | | : | | | : | | : |
QY 80 TNCARYTNTGYCAYGGNATNGARTAYCARAAAYATGMGNYTNCNAAYYTNTNGNCAYG 139
| | : | | | : | | | : | | : | | | : | | | : | | : |
Db 582 cgaaccaagaggacgcggcctcgagggtgcaccagttctaccctctgtgtaaaaggtgcagt 641
| | : | | | : | | | : | | : | | | : | | | : | | : |
QY 140 ARACNATGAARGARTNYTNGARCAGCNGGNCNTGGATNCCNTNGTNATGAARCART 199
| | : | | | : | | | : | | : | | | : | | | : | | : |
Db 642 gttctcctgagctacgcttctcttatgtctctatgtacgcaccgctgtgcaccg-tg--c 698
| | : | | | : | | | : | | : | | | : | | | : | | : |
QY 200 GYCAVCNGAYACNAARAARTTYTNTGYWSNYTNTTYGCNCCNGTNGYTTNGAYGAYY 259
| | : | | | : | | | : | | : | | | : | | | : | | : |
Db 699 tcgaccaagccattcctccgtgcggttccttgcgagcgcgcgcgacagggctgcgagg 758
| | : | | | : | | | : | | : | | | : | | | : | | : |
QY 260 TNGAYGARACNATNCARCCNTGYCAYWSNYTNTGYTNCARGTNAARGAYMGNTGYGCNC 319
| | : | | | : | | | : | | : | | | : | | | : | | : |
Db 759 cgctcatgaacaagttcggcttcacagtggccagagcggttgcgctgcgagaaactccc 816
| | : | | | : | | | : | | : | | | : | | | : | | : |
QY 320 CNGTNATGWSNGCNTTYGGNTTYCCNTGGCCNGAYATGYTNGARTGYGAYMGNTTYCC 377

RESULT 8 CEUC3557 1913 bp mRNA INV 09-OCT-1996
LOCUS Caenorhabditis elegans transmembrane receptor LIN-17 (lin-17) mRNA,
DEFINITION complete cds.
ACCESSION U63557
NID g1589779
KEYWORDS .
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans
Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;
Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;
Rhabditidae; Caenorhabditis.
REFERENCE 1 (bases 1 to 1913)
AUTHORS Sawa,H., Lobel,L. and Horvitz,H.R.
TITLE The Caenorhabditis elegans gene lin-17, which is required for
certain asymmetric cell divisions, encodes a putative
seven-transmembrane protein similar to the Drosophila frizzled
protein
Genes Dev. 10 (17), 2189-2197 (1996)
JOURNAL MEDLINE 96397419
REFERENCE 2 (bases 1 to 1913)
AUTHORS Sawa,H. and Horvitz,H.R.
TITLE Direct Submission
JOURNAL Submitted (11-JUL-1996) Biology, Massachusetts Institute of
Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA
FEATURES Location/Qualifiers
source 1..1913
/organism="Caenorhabditis elegans"
/strain="N2"


```

/organism="Homo sapiens"
/clone_lib="fz-2"
/tissue_type="ovary"
/chromosome="17"
/map="17q21.1"
133..1830
CDS
/gene="frizzled"
/function="development"
/note="polarity gene; putative"
/codon_start=1
/db_xref="PID:g736679"
/translation="MRPRSALPRLLLPLLLPAAGPAQFHGKEGISIPDHGFCQIPISI
PLCTDIAYNOTIMPNLLGHTNQEDAGLEVHQFPLVKVQCSPELRFELCSMYAPVCTV
LEQAIPPCRSICERARQGEALMNKFGFQWPERLRCHEHFPRHGAEOICVGONHSEDGA
PALLTTAPPGLQPGAGGTPGGPGGGAPPRYATLEHPFHCPRVLKVPSYLSYKFLGE
RDCAAPCEPARPDGSMFFSQEETRFARLWLTWSVLCCASTFFTTVTVLVDNQRFRRYP
ERPIIFLSGCTMVSVAYIAGFVLQERVVNERFSEDGYRTVVQGTKEGCTILFMMML
YFFSMASSIWWILSLTWFLAAGMKWGHEAIEANSQYFHLAAWAVPAVKTTILAMGQ
IDGDLISGVCVGLNSLDPLRGFVLAPLFVLYFGTSFLLAGFVSLFRIRTIMKHDGT
KTEKLERLMVRIGVFSVLYTPATIVIACYFYEQA FREHWERSWVSQHCKSLAIPCPA
HYTPRMSPDFTVYMIKYLMTLIVGITSGFWISGKTLHSWRKFYTRLTNSRHGETTV"
BASE COUNT      306 a      671 c      590 g      356 t
ORIGIN
```

```

Query Match      7.8%;      Score 54;      DB 80;      Length 1923;
Best Local Similarity 39.7%;      Pred. No. 8.23e-08;
Matches 119;      Conservative 41;      Mismatches 137;      Indels 3;      Gaps 2;

Db 263 tcccgctgtgcacggacatgcctacaacagaccatcatgcccaaccttctgggccaca 322
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 80 TNCARYTNTGYCAYGGNATNGARTAYCARAAAYATGMGNYTNCNAAYYTNTNGNCAYG 139
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 323 cgaaccaggaggacgcaggccctagaggtgcaccagttctatccgctggtagaggtgcagt 382
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 140 ARACNATGAARGARGTNYTNGARCARGCNGGNCNTGGATNCCNYTNGTNATGAARCART 199
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 383 gctcgccccgaactgcgcttctctgtgctccatgtacgcacccgctgtgcaccg-tg--c 439
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 200 GYCAIYCCNGAYACNAARAARTTYTNTGYWSNYTNTTYGCNCCNGTNTGYTNGAYGAYY 259
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 440 tggaaacaggccatccccgcgctgctctatctgtgagcgcgcgcagcggctgcgaag 499
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 260 TNGAYGARACNATNCARCCNTGYCAYWSNYTNTGYTNCARGTNAARGAYMGNTGYGCNC 319
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 500 ccctcatgaacaagtgcggttttcagtgcccgagcgcctgcgctgcgagcacttccgcg 559
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 320 CNGTNATGWSNGCNTTYGGNTTYCCNTGGCCNGAYATGYTNGARTGYGAYMGNTTYCCNC 379
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

RESULT 11
LOCUS      RATERZH      1912 bp      mRNA      ROD      16-JUL-1993
DEFINITION      Rattus norvegicus Drosophila polarity gene (frizzled) homologue
               mRNA, complete cds.
ACCESSION      L02530
NID            9310114
KEYWORDS      frizzled gene; homologue; polarity gene.
SOURCE      Rattus norvegicus (strain Sprague-Dawley) osteosarcoma cDNA to
               mRNA.
ORGANISM      Rattus norvegicus
               Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
               Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
               Murinae; Rattus.
REFERENCE      1 (bases 1 to 1912)
AUTHORS      Chan,S.D.H., Karpf,D.B., Fowlkes,M.E., Hooks,M., Bradley,M.S.,
               Vuong,V., Bambino,T., Liu,M.Y.C., Arnaud,C.D., Strewler,G.J. and
               Nissensohn,R.A.
TITLE      Two homologs of the Drosophila polarity gene frizzled (fz) are
               widely expressed in mammalian tissues
JOURNAL      J. Biol. Chem. 267, 25202-25207 (1992)
MEDLINE      93094228
FEATURES      Location/Qualifiers
               source
               1..1912
               /organism="Rattus norvegicus"
```

```

/strain="Sprague-Dawley"
/cell_line="UMR106-01"
/cell_type="osteoblast"
/tissue_type="osteosarcoma"
79..1791
CDS
/note="Drosophila polarity gene (frizzled) homologue"
/codon_start=1
/db_xref="PID:g310115"
/translation="MRARSALPRSALPRLLLPLLLPAAGPAQFHGKEGISIPDHGFC
QPISIPCLTDIAYNOTIMPNLLGHTNQEDAGLEVHQFYPLVKVQCSPELRFELCSMYA
PVCITVLEQAIPPCRSICERARQGEALMNKFGFQWPERLRCHEHFPRHGAEOICVGONH
SEDGTALLTTAPPSGLQPGAGGTPGGPGGGAPPRYATLEHPFHCPRVLKVPSYLSY
KFLGERDCAAPCEPARPDGSMFFSHHHTRFARLWLTWSVLCCASTFFTTTSLVAMQ
RFRYPERPFIIFLSGCTMVSVAYIAGFVLQERVVNERFSEDGYRTVGGTKKEGCTI
LFMMLYFFSMASSIWWILSLTWFLAAGMKWGHAIEANSQYFHLAAWAVPAVKTTI
LAMGQIDGDLISGVCVGLNRLDPLRGFVLAPLFVLYFGTSFLLAGFVSLFRITIM
KHDGTKTEPLERLMVRIGVFSVLYTPATIVIACYFYEQA FREHWERSWVSQHCKSLA
IPCPAHYTPRTSPDFTVYMIKYLMTLIVGITSGFWISGKTLHSWRKFYTRLTNSRHG
ETTV"
BASE COUNT      299 a      662 c      566 g      385 t
ORIGIN
```

```

Query Match      7.6%;      Score 52;      DB 90;      Length 1912;
Best Local Similarity 39.3%;      Pred. No. 5.83e-07;
Matches 118;      Conservative 41;      Mismatches 138;      Indels 3;      Gaps 2;

Db 224 tcccgctgtgcacggacatgcctacaacagaccatcatgcccaaccttctgggccaca 283
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 80 TNCARYTNTGYCAYGGNATNGARTAYCARAAAYATGMGNYTNCNAAYYTNTNGNCAYG 139
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 284 cgaaccaagaggacgcggccctggaggtgcatcaattctaccgctggtagaggtgcagt 343
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 140 ARACNATGAARGARGTNYTNGARCARGCNGGNCNTGGATNCCNYTNGTNATGAARCART 199
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 344 gctcgccccgagctgcgcttctctgtgctccatgtacgctccggtgtgcacgg-tg--c 400
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 200 GYCAIYCCNGAYACNAARAARTTYTNTGYWSNYTNTTYGCNCCNGTNTGYTNGAYGAYY 259
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 401 tggagcaggccatccccgcgctgcgctccatctgcgaacgcgcgcgaaggctgcgagg 460
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 260 TNGAYGARACNATNCARCCNTGYCAYWSNYTNTGYTNCARGTNAARGAYMGNTGYGCNC 319
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 461 cgctcatgaacaagttcggcttccagtgcccgagcgcctcgcgtgcgagcatttccgcg 520
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 320 CNGTNATGWSNGCNTTYGGNTTYCCNTGGCCNGAYATGYTNGARTGYGAYMGNTTYCCNC 379
| | :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

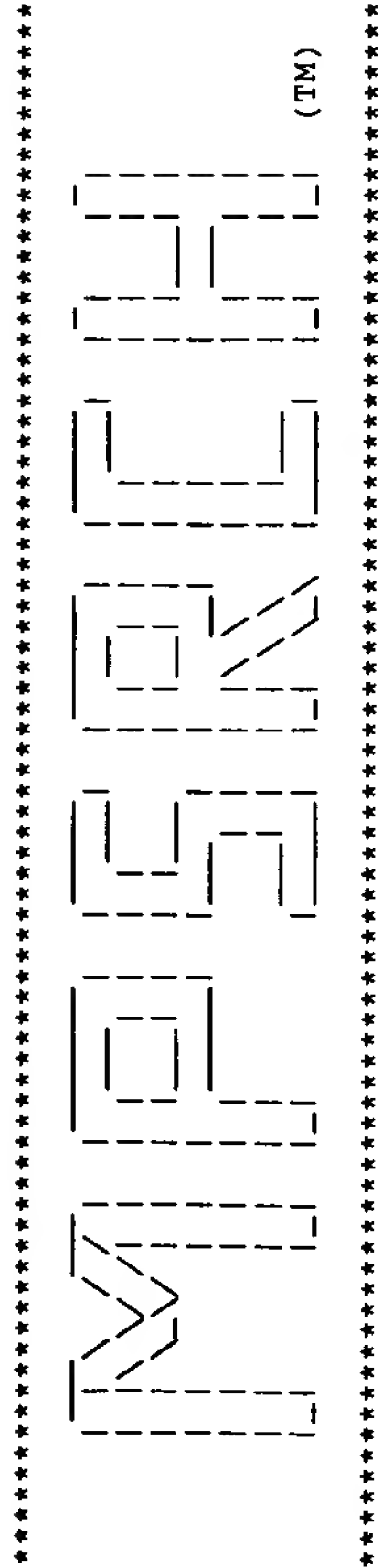
RESULT 12
ID      CET23D8      standard;      DNA;      INV;      34576 BP.
AC      281128;
NI      e1008304
DT      21-OCT-1996 (Rel. 49, Created)
DT      01-MAR-1997 (Rel. 51, Last updated, Version 5)
DE      Caenorhabditis elegans cosmid T23D8
KW      CD63 antigen like; Drosophila tissue polarity protein like;
KW      histone H2A; histone H4; Mouse bright protein like;
KW      transporter protein.
OS      Caenorhabditis elegans
OC      Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;
OC      Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;
OC      Rhabditidae; Peloderinae; Caenorhabditis.
RN      [1]
RP      1-34576
RA      Wild A.;
RT      ;
RL      Submitted (21-OCT-1996) to the EMBL/GenBank/DBJ databases.
RL      Nematode Sequencing Project, Sanger Centre, Hinxton, Cambridge CB10
RL      1RQ, England and Department of Genetics, Washington University, St.
RL      Louis, MO 63110, USA. E-mail: jes@sanger.ac.uk or
RL      rw@nematode.wustl.edu
RN      [2]
RX      MEDLINE; 94150718.
RA      Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
```

RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J.,
RA Coulson A., Craxton M., Dear S., Du Z., Durbin R., Favello A.,
RA Fulton L., Gardner A., Green S., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., McMurray A., Mortimore B.,
RA O'Callaghan M., Parsons J., Percy C., Rifken L., Roopra A.,
RA Saunders D., Shownkeen R., Smaldon N., Smith A., Sonhammer E.,
RA Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaudin M.,
RA Vaughan K., Waterston R., Watson A., Weinstock L.,
RA Wilkinson-Sproat J., Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of
RT C. elegans";
RL Nature 368:32-38(1994).
CC Current sequence finishing criteria for the C. elegans genome
CC sequencing consortium are that all bases are either sequenced
CC unambiguously on both strands, or on a single strand with both
CC a dye primer and dye terminator reaction, from distinct
CC subclones. Exceptions are indicated by an explicit note.
CC
CC IMPORTANT: This sequence is NOT necessarily the entire insert
CC of clone T23D8. It may be shorter because we only sequence
CC overlapping sections once, or longer because we arrange for a
CC small overlap between neighbouring submissions.
CC The true left end of clone T23D8 is at 1 in this sequence.
CC The true right end of clone T23D8 is at 34576 in this sequence.
CC The true right end of clone T24D1 is at 12749 in this sequence.
CC Coding sequences below are predicted from computer analysis,
CC using the program Genefinder (P. Green, ms in preparation),
CC and other available information.
CC The start of this sequence (1..104) overlaps with the end of
CC sequence CER24D1.
FH Location/Qualifiers
FH
FH source 1..34576
FT /organism="Caenorhabditis elegans"
FT /clone="T23D8"
FT /chromosome="I"
FT complement(join(31876..31973,32210..32284,32331..32573,
FT 32640..33032,33192..33264))
FT /product="T23D8.8"
FT /note="protein predicted using Genefinder"
FT /note="Similarity to Mouse bright gene (TR:Q62431)"
FT /db_xref="PID:e304447"
FT /translation="MYELSDDVKRKEWLDLWLFMHRIKGPVTRIPIMAKQVLDLYEL
Y
FT RLVVQHGLVEIINKKLWREITKGLNLPSSITSAFLTQYQKYLYDYECEKEKLSN
Q
FT SDLQQAIDGNRRREAPGRTAPSPFLPFQLPHAASAAATMLNNQLGLMRNDLDDEN
T
FT LSLQASGLFGTSYGAEQMAILEAHQHNLERAQRAVOQQVARQSLGLTACSNNGGNIH
N
FT SGRESTSSNDSIDIPAKRPKLENDVKTNGASSMRISTKHSDNSKTSMSVSMEINGITYQ
G
FT VLFALDETVSES"
FT complement(25597..25908)
FT /product="T23D8.5"
FT /note="protein predicted using Genefinder"
FT /note="Similarity to histone H4"
FT /db_xref="PID:e275927"
FT /translation="MSGRGKGKGLGKGAKRHRKVLRDNIQGITKPAIRRLARRGGV
K
FT RISGLIYEETRGVLKFLENVIRDVATYCEHAKRKTVTAMDVVYALKRQGRITLYGFGG
T
FT complement(join(7258..7391,7494..7638,8798..8860,9211..934
3
FT ,10019..10101,10150..10212))
FT /product="T23D8.2"
FT /note="protein predicted using Genefinder"
FT /note="Similarity to Rat CD63 antigen (SW:P28648)"
FT /db_xref="PID:e276248"
FT /translation="MVEGGVTIVKYLLFLANLVLVWVGLSLIIVGSILQLKFDNVLDI
L
FT GDERLATPILLLVIGSLCTLLGLGCCGAIRENYCLTVSFAVLLALLITGMVRYHESR

G
FT VESAWDKTHQLFECCGVTNSSDWLTFTTIPDSCCIEIEGCARENAPLFEPCGHSVE
Q
FT WVLKNGAMVGGICAVLAAIQLVGVCFACCLSKSILKDFHDFYY"
FT join(2786..3258,3758..4122,4187..4325,4866..5107,5220..534
O
FT ,5688..5906,6332..6479)
FT /product="T23D8.1"
FT /note="protein predicted using Genefinder"
FT /note="Similarity to Drosophila Tissue polarity protein
FT (TR:Q24760)"
FT /db_xref="PID:e275930"
FT /translation="MHRHILILFEGCLSDQRLSSTSISMMNGFSTTRKCEHITIPM
C
FT KNLDYNQTVFPNLLGHTTQSEAGPAIAQFNPLIKVKCEDIRLFLCTVYAPVCTVLEK
P
FT IQPCRELCLSAKNGCESLMKKFGQWPDQLDCNKFPVTDLCVGKNSESSENKKSSND
V
FT TFGVSTIANEVVLSPPKKCPHHMHTTSGSHFSLPLLSGRLPECSLTCEADNQVPMFDFG
R
FT VRRILRIWTAASVACFVCSLFTLVTLVDLSRFAYVVRPILYLAFCYLAISTVYMIG
V
FT VGEDGFACGTGYGSTPTTLVTQGGENVGCSALAVVHYFFFMSSCAWWLVLCLAWFLAAN
L
FT KWGAESIAALSIFYHAMCWGVPVAVLSVTVLVTVNSVDGVDFTGICSVGNLNPALVYFF
F
FT TPIVVSLALGAVLLVCGIWSMIRIRSIYIKLQHADVERNISKLEKLMRLIGAFAIMYSL
P
FT TAMNAAIMWYQAVNMPAWLEGWLHRCVRLQDRELFGTYPVDDCPMDPKVAAPEIIV
F
FT LLKYVSQLVVGITCAIWWVSSKTLSSYHKAYLALSSRSPTVPAHVDQVNMR"
FT 26277..26660
FT /product="T23D8.6"
FT /note="protein predicted using Genefinder"
FT /note="Similarity to histone H2A"
FT /note="CDNA EST yk72a2.3 comes from this gene"
FT /note="CDNA EST yk124d7.3 comes from this gene"
FT /note="CDNA EST yk124d7.5 comes from this gene"
FT /note="CDNA EST yk72a2.5 comes from this gene"
FT /db_xref="PID:e276250"
FT /translation="MSGRGKGKAKTGKAKSRSSRAGLQFPVGRHLRILRKGNYAQR
V
FT GAGAPVYLAAVLEYLAAEYVLELAGNAARDNKKTRIAPRHLQLAVRNDDELNKLLAGVT
I
FT AOGGVLPNIOAVLLPKKTGGDKE"
FT join(26949..27068,27120..27187,27232..27606,27661..28621,
FT 28689..28819,29049..29748,30273..30464,30510..30659)
FT /product="T23D8.7"
FT /note="protein predicted using Genefinder"
FT /note="CDNA EST yk116f3.3 comes from this gene"
FT /note="CDNA EST yk20f1.5 comes from this gene"
FT /note="CDNA EST yk116f3.5 comes from this gene"
FT /db_xref="PID:e275933"
FT /translation="MEDQWLLSAIYDDDLVEKLVKRSSTSSRSTINVPSENEFLSS
S
FT SGRVSDDLYLHPHIEENREPPFKLIGKPLPSTTGRFLSLLANHFQITCNGSIIHQYIIR
F
FT DPDIPSKKLNRTILRTLQEQNPGLIECPLVFDGIHTVYSTELINVKEVNSVINVAGV
V
FT NTKESPNLFKLYLTHVDSFLDDTKIITGNQDQNKLRMMHAIDTVFRQTSTGNFHAVL
Q
FT SPFSAQNSAIEPSHGLGWGTVNLGVGREVCYGFYQNVVETFDTLTMNLDVATTTFYR
P
FT VALVEFLAEILEVPLATVTDGRSLSDVQKKKFNREVAGLKVETRHCSCPRRFRVARCT
W
FT KPTENISFHLSETAGNQDSKPLSLVEYYKRRYNIDLTYKHLPCIEVGRTRRECILPLEL
C
FT YVVGSGQRICKKLEQQIANLIRATSRNATERQNAVMSLQNLRLKMDNDVNAVFKGLKVE
A
FT QLLKIEGRVLPVPRLLYRSPNLKRQECVTVPNNGTWDMRGKNFYSGIQIREWAIVCFA
S
FT PEIIGEASMRSFVRNLNVVASEIGMPFLEEHRFCRYAEPDQTVKLLLEHLNEQYNLQLV
L

SOURCE ORGANISM	H2A; histone H4; Mouse bright protein like; transporter protein. Caenorhabditis elegans.
REFERENCE	Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
AUTHORS	1 (bases 1 to 34576)
TITLE	Wild,A.
JOURNAL	Direct Submission
REFERENCE	Submitted (21-OCT-1996) Nematode Sequencing Project, Sanger Centre, Hinxton, Cambridge CB10 1RQ, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail: jes@sanger.ac.uk or rwenematode.wustl.edu
AUTHORS	2 (bases 1 to 34576)
REFERENCE	Wilson,R., Ainscough,R., Anderson,K., Baynes,C., Berks,M., Bonfield,J., Burton,J., Connell,M., Copsey,T., Cooper,J., Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A., Fulton,L., Gardner,A., Green,P., Hawkins,T., Hillier,L., Jier,M., Johnston,L., Jones,M., Kershaw,J., Kirsten,J., Laister,N., Latreille,P., Lightning,J., Lloyd,C., McMurray,A., Mortimore,B., O'Callaghan,M., Parsons,J., Percy,C., Rifken,A., Roopra,A., Saunders,D., Showkeen,R., Smaldon,N., Smith,A., Sonhammer,E., Staden,R., Sulston,J., Thierry-Mieg,J., Thomas,K., Vaudin,M., Vaughan,K., Waterston,R., Watson,A., Weinstock,L., Wilkinson-Sproat,J. and Wohlman,P.
TITLE	2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans
JOURNAL MEDLINE	Nature 368 (6466), 32-38 (1994)
COMMENT	94150718 Current sequence finishing criteria for the C. elegans genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note. IMPORTANT: This sequence is NOT necessarily the entire insert of clone T23D8. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions. The true left end of clone T23D8 is at 1 in this sequence. The true right end of clone T23D8 is at 34576 in this sequence. The true right end of clone T24D1 is at 12749 in this sequence. Coding sequences below are predicted from computer analysis, using the program Genefinder (P. Green, ms in preparation), and other available information. The start of this sequence (1..104) overlaps with the end of sequence CET24D1.
FEATURES source	Location/Qualifiers 1..34576 /organism="Caenorhabditis elegans" /clone="T23D8" /chromosome="I" join(2786..3258,3758..4122,4187..4325,4866..5107,5220..5340,5688..5906,6332..6479) /note="protein predicted using Genefinder; Similarity to Drosophila tissue polarity protein (TR:Q24760)" /codon_start=1 /product="T23D8.1" /db_xref="PID:e275930" /db_xref="PID:g1628244" /translation="MHRHILILFLGCLSDQRLSSTSISSMNGFSTTRKCEHITIPMCKNLDYNQTVFNLGHTQSEAGPAIAQFNPLIKVKCEDIRLFLCTVYAPVCVLEKPIQPCRELCSAKNGCSLMKKFGFWPDQLDCNKFPTDLCVGNKSSSNRSKSSNDVTFGVSTIANEVLSPKCPHHMHTTSGSHFSLPLSGRLPECSLTCEADNQVMMFDGVRRIILRIWTAAWSVACFVCSLFTLVTLFVLDLSRAYPVRPILYLAFCYLAISTVYMGVGEDGFACGTGCTPTTLVTQGGENVGCSALAVVHYFFFMSSCAWWLVLCIAWFLAANLKGAESIAALSPYFHAMCWGPVAVLSVTLVNTNSVDGDFVTGICSVGNLNSALVYFFFTPIVVSALGAVLLVCGIWSMIRIRSYIKQHADVERNISKLEKLMRIGAFAIMYSLPTAMNAAMWYQAVNMPAWLEGWLHRCVRLQDRELFGFTYFVDDCPMDPKVAAPAEIIVLLKYVSQLVVGITCAIWWVSSKTLSSYHKAYLALSSRSPTVPAHVDQVNR"
CDS	complement(join(7258..7391,7494..7638,8798..8860,9211..9343,10019..10101,10150..10212)) /note="protein predicted using Genefinder; Similarity to

Rat CD63 antigen (SW:P28648)"	CDS
/codon_start=1	
/product="T23D8.2"	
/db_xref="PID:e276248"	
/db_xref="PID:g1628243"	
/translation="MVEGGVTIVKYLFLANLVWVGISLIIVGSILQLKFDNVLDIIGDERLATPILLVIGSLCTLLGFLGCCGAIRENYCLTVFAVLLALLITGMVRYHESRGVESAWDKTHQLFECGVTNSSDWLTFTTIPDSCCIEEGCARENAPLFEPCGCIHSVEQVLKNGAMVGGICAVLAATQLVGVCFACCLSKSLKDFHDFY"	
complement(join(10723..10767,11212..11658,11707..11823,12477..12917,13249..13431))	
/note="protein predicted using Genefinder; cDNA EST cm16f9 comes from this gene; cDNA EST CEESF93R comes from this gene; cDNA EST CEESF93FC comes from this gene"	
/codon_start=1	
/product="T23D8.3"	
/db_xref="PID:e304448"	
/db_xref="PID:g1865750"	
/translation="MGKKRAFIDKKASQSFRLVPDNRRSERSERFKPTQEHLEEQQKYGYYDDDYDYLQHMRAVNEPMKLENVHEEVEKTTIKTSGSAFPPAPPPLFGLVGALKKPEFFDEVDANALEEVTDDRTNGELEDNFITLAGGLDERTVYRSTRGEDSEEEEDDDDEDEMYDDYNDDDELFGEEAVGEIRVERADQVIDNAFEELMDREYNTDQIGELDGGDDYVGGALEPNAGRLHKLINDKGPSNAEYDELAKHVYVRMRMLIEEGVIKDKEEYIEIVDEGTNKKMKWDCSEFATQYNIYNHPTLIKEPRGLSRKALKRFDKAVEEMDIAEEDDDDEDDDDDDKESVSTVSTFRPKNETPEQSRSLRKKAVERKLRREKANKTMFAE	
EKRKLAKGRIGQIKARPI"	
complement(join(13883..14240,14292..15409,15456..16341,16423..16656,16705..16807,17668..17721,17774..18278,18330..18656,18888..19113,19170..19705,19751..19828,19884..19991,20041..20214))	
/note="Similarity to Yeast nuclear transport protein (SW:P32497); cDNA EST yk44d4.5 comes from this gene; cDNA EST yk33b1.5 comes from this gene; cDNA EST yk37d3.5 comes from this gene; cDNA EST yk8e5.5 comes from this gene; cDNA EST yk8e7.5 comes from this gene; cDNA EST yk4b7.5 comes from this gene; cDNA EST yk30g3.5 comes from this gene; cDNA EST yk8e5.3 comes from this gene; cDNA EST yk8e7.3 comes from this gene; cDNA EST yk44d4.3 comes from this gene; cDNA EST yk37b1.3 comes from this gene; cDNA EST yk37d3.3 comes from this gene"	
/codon_start=1	
/product="T23D8.4"	
/db_xref="PID:e304461"	
/db_xref="PID:g1865751"	
/translation="MSYHOGDDRRMSMLSVNTTMTNQFPDSQRCYSSNGSTCENVNPEMQMNIQTAAEQAIRLWFNTFDPMQRLHMAKTIRTWIRODKFAQVDQANMPCVQQILNIYDGLKQPQVQLPISYYAQLWYNLLDILRRFTPIISPIYHQVVMFCPRENGPQDFRELICNLISLNWQKDPHKHMCANQVQIFNCIIMGVYNEKRLTEFAQHLKFEKLVGTLSSEYFNQVHPGMINPAIFIIFRFIISKDTRLKDYFIWNNNPHDQPPPTGLIIKLNAVMIGSYRLIAGQNPETLPQNPDLAHLIQVIIRTFDGLLLHSDAIDGFVRSDBGVGAITTVVQYPNNDLIRAGCKLLQVSDAKALAKTPLENIPLFLLRIEIHDPDEVIYSGTGFLSNVVAHKQHVKDIAIRSNALFLLHTIISKYPRDELTDAPKRNRVCEIICNCLRTLNFLMWIPTPNGETKTAGPNEKQVCKFIEIDILKMLSCLSCEGMDTPGLLELRS	
TILRSFILLRLRTPFPVKDGYLVNIDENRKENLGHICAAYSWVFRQPNNTRTQSTKQQLVERTISLLVLMEQCGAEVAQYSYSDCPLNLLNGQVKPTFIHNVLVVCDKILEHCPTRADIWTIDRPMLEGLTNHRNSDIAKAANSLLSRFPENDLLAAMSFFFHAKEDSD	
SDTSSSEDEVEDQKVNSAKFRDDLDFMAGPEEDEKRVVRAQDKKFDKELGIKQNRDAKSNKDLNRLLTGTDSLAKAYDKSTVFQRONVANPRFYIRSLVEIDYVGNKLWDDK	
DAKSALSKNNAKALPLRQKLKYYIKDQQLDLDVTRYNPDGEGYETPEDEDDDDFG	
EVSESKAEKSPGKPSEKAASDSDSDSDSDSDSSNWSSEPESSDDEDSVTKMEQLRR	
YFLKKEFRVESKDDKKEKKRVIRVKEAVEEDDDADDPVYNREKSVVHFDPKKEEVTHD	
VMKKLNEVMSARGKRTTDRNQHVANLRKLLLEVSEKEKLGILGINISFCIISALFEL	
NAKISDHMEYETFTTLOFVNSLLDLLIGTDVRKLSVTAEDENLKDQDTQEIYIQGS	
ILIAVQRLDGELAKILQNADCHSNDYIEKLAEKDMCSLIEKAEKYVELRNSDGIFDK	
HEVCKVYMMRIEHAYYKYQDQNEEDAGKLDYLCNKIYTLDDKRLRQRAMLCHVYYL	
AVHDKWHRARDLLMSHMOAIVDHSVDVDTQILNRTICQLGCAFRHGFIREAHQGLS	
EIQNTQRAKELLAAQAVGTQHEKTAEQEKIDRSRQVPYHMHINVELMECVYLICSMLL	
EIPHMASCEFEMRRRMLSRPHYQLKQSEKASLTGPPNREHVVAASKAMLNGDWKK	
CQDYIVNDKMNQVWNLFHNAETVKGMVVRRIQEEESLRYLLTYSTVATVSLKKLAD	
LFELSKDKDVHSIISKMIIOEELSATLDEPTDCLIMHRVPSRLQMLALNLSDKLQTLA	
ENNEQILEPRTGRGGYQGPGSWFFPGRNERQGDQKQKSGGYQGERRGQGDGKRGNGW	
SQGGQQRHRHPQKPRAF"	



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

Mpsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 9 08:43:36 1997; MasPar time 1043.08 Seconds
Tabular output not generated. 997.528 Million cell updates/sec

Title: >US-08-848-439-1
Description: (1-2027) from US08848439.seq
Perfect Score: 2027
N.A. Sequence: 1 GAATTCGGCCTTCATGGCCT.....AAAAAAGCGGCGC 2027
Comp: CTTAAGCCGGAAGTACCGA.....TTTTTTTTCGCGGCG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 707517 seqs, 256659390 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

EST-STS	
1:EST1	2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9	10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15	16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21	22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27	28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33	34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39	40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45	46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51	52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57	58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63	64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69	70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75	76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81	82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87	88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
93:EST93	94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
99:EST99	
EST-STS-TWO	
100:EST100	101:EST101 102:EST102 103:EST103 104:EST104
105:EST105	106:EST106 107:EST107 108:EST108 109:EST109
110:EST110	111:EST111 112:EST112 113:EST113 114:EST114
115:EST115	116:EST116 117:EST117 118:EST118 119:EST119
120:EST120	121:EST121 122:EST122 123:EST123 124:EST124
125:EST125	126:EST126 127:EST127 128:EST128 129:EST129
130:EST130	131:EST131 132:EST132 133:EST133 134:EST134
135:EST135	136:EST136 137:EST137 138:EST138 139:EST139
140:EST140	141:EST141 142:EST142 143:EST143 144:EST144
145:EST145	146:EST146 147:EST147 148:EST148 149:EST149
150:EST150	151:EST151 152:EST152 153:EST153 154:EST154
155:EST155	156:EST156 157:EST157 158:EST158 159:EST159
160:EST160	161:EST161 162:EST162 163:EST163 164:EST164
165:EST165	166:EST166 167:EST167 168:EST168 169:EST169
170:EST170	171:EST171 172:EST172 173:EST173 174:EST174

175:EST175 176:EST176 177:EST177 178:EST178 179:EST179
180:EST180 181:EST181 182:EST182 183:EST183 184:EST184
185:EST185 186:EST186 187:EST187 188:EST188 189:EST189
190:EST190 191:EST191 192:EST192 193:EST193 194:EST194
195:EST195 196:EST196 197:EST197 198:EST198

Statistics: Mean 12.162; Variance 2.824; scale 4.307

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result NO.	Score	Query Match	Length DB ID	Description
1	473	23.3	480 134	yy64d09.r1 Homo sapie
2	453	22.3	460 173	zd70ell.s1 Soares fet
3	416	20.5	446 101	yy64d09.s1 Homo sapie
4	413	20.4	462 174	zd70ell.r1 Soares fet
5	408	20.1	445 92	yw83c06.s1 Homo sapie
6	363	17.9	433 156	ze76gl0.s1 Soares fet
7	359	17.7	426 72	ys74d07.r1 Homo sapie
8	337	16.6	406 61	yl28c05.r1 Homo sapie
9	331	16.3	340 48	Human aorta CDNA 5'-e
10	328	16.2	396 178	mm97f07.r1 Stratagene
11	329	16.2	414 156	ze76gl0.r1 Soares fet
12	322	15.9	434 46	yo73a07.r1 Homo sapie
13	321	15.8	469 61	yl28c05.s1 Homo sapie
14	319	15.7	326 92	yw83e07.s1 Homo sapie
15	316	15.6	330 48	Human aorta CDNA 5'-e
16	313	15.4	344 99	yw83c06.r1 Homo sapie
17	299	14.8	301 110	Human aorta CDNA 5'-e
18	299	14.8	313 48	Human aorta CDNA 5'-e
19	298	14.7	303 49	Human aorta CDNA 5'-e
20	294	14.5	300 48	Human aorta CDNA 5'-e
21	290	14.3	293 48	Human aorta CDNA 5'-e
22	290	14.3	293 48	Human aorta CDNA 5'-e
23	290	14.3	344 61	yl26d03.r1 Homo sapie
24	287	14.2	299 48	Human aorta CDNA 5'-e
25	285	14.1	287 48	Human aorta CDNA 5'-e
26	279	13.8	281 110	Human aorta CDNA 5'-e
27	273	13.5	273 48	Human aorta CDNA 5'-e
28	269	13.3	283 164	Human aorta CDNA 5'-e
29	256	12.6	263 110	Human aorta CDNA 5'-e
30	254	12.5	266 48	Human aorta CDNA 5'-e
31	251	12.4	300 49	Human aorta CDNA 5'-e
32	241	11.9	241 48	Human aorta CDNA 5'-e
33	230	11.3	247 110	Human aorta CDNA 5'-e
34	226	11.1	227 61	yl26d03.s1 Homo sapie
35	213	10.5	215 164	Human aorta CDNA 5'-e
36	212	10.5	222 73	Human aorta CDNA 5'-e
37	210	10.4	216 48	Human aorta CDNA 5'-e
38	209	10.3	299 148	mb41f02.r1 Soares mou
39	200	9.9	201 140	HUMGS0005382, Human G
40	200	9.9	231 65	ym31e03.r1 Homo sapie
41	200	9.9	258 98	JJ9816F Homo sapiens
42	194	9.6	198 182	15684.seq.F Fetal hea
43	193	9.5	195 182	cl0685.seq.F Fetal he
44	192	9.5	204 65	ym31e03.s1 Homo sapie
45	184	9.1	210 48	Human aorta CDNA 5'-e

ALIGNMENTS				
RESULT 1	N93894	480 bp	EST	05-APR-1996
LOCUS	yy64d09.r1 Homo sapiens	CDNA clone 278321 5'		
DEFINITION	N93894			
ACCESSION	g1266203			
NID	EST.			
KEYWORDS	human clone=278321 primer=mb.REGA+ET library=Soares multiple			
SOURCE				

Db 61 tatttttacaactttttaeaaggtaaaactactatgtatatcacaggtgaagctacaatgggtt 120
|||||
Cp 1926 TATTTTACAACCTTTTAAAGGTAAACTACTATGTATATACAGGTAAGCTACAATGGGTT 1867
|||||
Db 121 taatttgcaaaagttaagtaagaagaatgtttttaaacaaggcttaaaagtactcaagtcaatt 180
|||||
Cp 1866 TAATTGCAAAAGTTAAGTAAGAAATGTTTAAACAAGGCTTAAAGTACTCAAGTCAATT 1807
|||||
Db 181 ataaaatttatatcttttgccttttacttgaagaaatcatcgctatagaaaatgggttaatgt 240
|||||
Cp 1806 ATAAATTTATATCTTTTGCCTTTTACTTTGAAGAAATCATGCTATAGAAATGGTTAATGT 1747
|||||
Db 241 gcttctaataaatggaagtattgttagctggaatgtgatacatgtgaacagtttaagttccc 300
|||||
Cp 1746 GCTTCTAATAATGGAAGTATTGTAGCTGGAATGTGATACATGTAAACAGTTTAAAGTCCC 1687
|||||
Db 301 attgaaggtataaaatgatgaattcgttgaagacttagacactgagtcctcagtcctggag 360
|||||
Cp 1686 ATTGAAGGTATATAAATGATGAATT-GTTGTAAGACITAGACACTGAGTCTCAGTCTGGAG 1628
|||||
Db 361 ctgatgaagatgttgagataaacagccagcttttatctcaacagggtttgtgaccccaagt 420
|||||
Cp 1627 CTGATGAAGATGTTGAGATAACAGCCAGCTTTTATCTCAACAGGGTTTGTGACCCCAAGT 1568
|||||
Db 421 ttggggccacagagaaaaattgaagcaatttgcatgttatga 460
|||||
Cp 1567 TTGGGCCACAGAGAAAAATTGAAGCAATTGTCATGTTATGA 1528
|||||

RESULT 3
LOCUS N64004 446 bp mRNA EST 01-MAR-1996
DEFINITION yy64d09.s1 Homo sapiens cDNA clone 278321 3'.
ACCESSION N64004
NID 91211833
KEYWORDS EST.
SOURCE human clone=278321 primer=ml3 -40 forward library=Soares multiple sclerosis 2NBHMSF vector=PT7T3D (Pharmacia) with a modified polylinker V-type: phagemid host=DH10B (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI 46 year old male. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGATTTTTTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from 4 multiple sclerosis lesions from one patient was kindly provided by Dr. Kevin G. Becker (NINDS/NIH).

ORGANISM Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 446)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
REFERENCE 1 (bases 1 to 446)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 225
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES
source Location/Qualifiers
1..446
/organism="Homo sapiens"
/clone="278321"
mRNA
BASE COUNT 157 a 51 c 87 g 149 t 2 others
ORIGIN
Query Match 20.5%; Score 416; DB 101; Length 446;
Best Local Similarity 98.9%; Pred. No. 0.00e+00;
Matches 440; Conservative 0; Mismatches 1; Indels 4; Gaps 4;
Db 2 taggtaaaaacaggatgtaaaagtttatatacaagaataataatgtttatctgaaaatatattac 61
|||||
Cp 2003 TAGGTAAAAACAGGATGTAAGTTTATATACAAAGATATAATGTTTATCTGAAATATTAC 1944
|||||
Db 62 egtgttggttaaagcaatattttttacaacttttaaaggtaaactactatgtatatattacag 121
|||||
Cp 1943 AGTGTGTGTTAAAGCAATATTTTTTACAACCTTTTAAAGGTAAACTACTATGTATATTACAG 1884
|||||
Db 122 gtaagctacaatggggtttaatttgcaaaaagttaaagtaagaaaatgttttaaacaaaggctta 181
|||||
Cp 1883 GTAAGCTACAATGGGTTTAAATTGCAAAAAGTTAAGTAAGAAATGTTTTTAAACAAGGCTTA 1824
|||||
Db 182 aagtactcaagtcaattataaaaatttatatcttttgccttttacttgaagaaatcatgct 241
|||||
Cp 1823 AAGTACTCAAGTCAATTATATAAAATTTATATCTTTTGCCTTTTACTTGAAGAAATCATGCT 1764
|||||
Db 242 atagaaatggttaatgtgcttctaataaaaatggaagtattgttagctgggaatgtgatacatg 301
|||||
Cp 1763 ATAGAAATGGTTAATGTGCTTCTTAATAAAATGAAGTATTGTAGCTGGAATGTGATACATG 1704
|||||
Db 302 taacagtttaagttcccatggaagggtataaaaatngatgaattgttgtaagaccttagaca 361
|||||
Cp 1703 TAACAGTTTAAAGTTCCCATTTGAAGGTATAAAAT-GATGAATTTGTTAAGAC-TTAGACA 1646
|||||
Db 362 ctgagtcctcagtcctgggagctgataagaatgttgagataaacgccagcttttatctcaac 421
|||||
Cp 1645 CTG-AGTCTCAGTCTGG-AGCTGATGAAGATGTTGAGATAAACGCCAGCTTTATCTCAAC 1588
|||||

Db 422 agggttgtgaccncaaagtttggg 446
|||||
Cp 1587 AGGGTTGTGACCCACCAAGTTTGGG 1563
|||||

RESULT 4
LOCUS W77968 462 bp mRNA EST 17-OCT-1996
DEFINITION zd70e11.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 346028 5'.
ACCESSION W77968
NID g1388502
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 462)
REFERENCE 1 (bases 1 to 462)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1550 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 424.
Location/Qualifiers

FEATURES
source

1..462
/organism="Homo sapiens"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCATCTTTTTTT-3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NbHL19W."
/clone="346028"
/clone_lib="Soares fetal heart NbHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
<1..>462

BASE COUNT 128 a 133 c 106 g 91 t 4 others
ORIGIN

Query Match 20.4%; Score 413; DB 174; Length 462;
Best Local Similarity 98.0%; Pred. No. 0.00e+00;
Matches 449; Conservative 0; Mismatches 3; Indels 6; Gaps 6;

Db 6 tcgctcttgcggccgctgcctcgatgacctagacgagaccatccagccatgccactcg 65
|||||
QY 544 TCGCTCTTGCggcccgctgcctcgatgacctagacgagaccatccagccatgccactcg 603
Db 66 ctctgn-tgcaggtgaaggaccgctgcggcccggtcatgtccgccttcgn-ttccccctgg 123
|||||
QY 604 CTCGTGCGTGCAGGTGAAGGACCGCTGCGCCCGGTCTATGTCCGCCCTTCGGCTTCCCTGG 663
Db 124 ccgacatgcttgagtcgaccgctttccccccaggacaacgacctttgcacccccctccgc 183
|||||
QY 664 CCGGACATGCTTGAGTGGACCGCTTCCCCCAGGACAACGACCTTTGCGATCCCCCTC-GC 722
Db 184 tagcagn-accacctctgccagccaccgagggaagctccaaaggatgtgaagcctgcaa 242
|||||
QY 723 TAGACGCGACCACTCTCTGCCAGCCACCGAGGAAGCTCCAAGGATGTGAAGCCCTGCAA 782
Db 243 aaataaaaaatgatgatgacaaacgacataatggaaacgctttgtaaaaatgatttgcact 302
|||||
QY 783 AAATAAAAAATGATGATGACAACGACATAATGGAAACGCTTTGTGAAAATGATTTGCACT 842
Db 303 gaaaaataaagtgaaggagataaacctacatcaaccgagataccaaaaatcatcctggaaga 362
|||||
QY 843 GAAAATAAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAAATCATCCTGGA-GA 901
Db 363 ccaagagcaagaccatttacaaagctgaacggtgtgtccgaaaaggaccttgaagaaaatcg 422
|||||
QY 902 CCAAGAGCAAGACCACTTTACAAGCTGAACGGTGTGTCCGAAAGGGACCT-GAAGAAATCG 960
Db 423 gtgctgtggctcaaagacagcttgcagtgacacctgtga 460
|||||
QY 961 GTGCTGTGGCTCAAAGACAGCTTTCAGTGCACCTGTGA 998

RESULT 5
LOCUS N32415 445 bp mRNA EST 10-JAN-1996
DEFINITION Yw83c06.s1 Homo sapiens cDNA clone 258826 3'.
ACCESSION N32415
NID g1152814
KEYWORDS EST.
SOURCE human clone=258826 primer=m13 -40 forward library=Soares placenta
8to9weeks 2NbHP8to9W vector=pT7T3D (Pharmacia) with a modified

polylinker host=DH10B (ampicillin resistant) Rsite1=Not I
Rsite2=Eco RI two placentae: one from 8 weeks and another from 9
weeks post conception. 1st strand cDNA was primed with a Not I -
oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCATCTTTT-3']
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT7T3 vector (Pharmacia). Library
constructed by Bento Soares and M. Fatima Bonaldo.

ORGANISM

Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 445)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)

TITLE
JOURNAL
COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
High quality sequence stops: 314

Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES

Location/Qualifiers
source
1..445
/organism="Homo sapiens"
/clone="258826"

BASE COUNT 156 a 55 c 87 g 147 t
ORIGIN

Query Match 20.1%; Score 408; DB 92; Length 445;
Best Local Similarity 98.6%; Pred. No. 0.00e+00;
Matches 434; Conservative 0; Mismatches 2; Indels 4; Gaps 4;

Db 1 aaaacaggatgtaaagtttatatacaagaataatgtttatctgaaatattacagtgt 60
|||||
Cp 1998 AAAACAGGATGTAAAGTTTATATACAGAATAATAATGTTTATCTGAAATATTACAGTGT 1939
Db 61 tggtaaagcaataattttacaacttttaaggtaaactactatgtatattacaggtaag 120
|||||
Cp 1938 TGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAAACTACTATGTATATTACAGGTAAG 1879
Db 121 ctacaatgggtttaatttgcaaaagttaagtaagaaatgttttaacaagaagccttaagta 180
|||||
Cp 1878 CTACAATGGGTTTAATTTGCAAAAGTTAAGTAAGAAATGTTTAAACAAGGCTTAAAGTA 1819
Db 181 ctcaagtcaattataaaaaatttatatcttttgccttttacttgaaagaaatcatgctataga 240
|||||
Cp 1818 CTCAAGTCAATTATAAAATTTATATCTTTTGCTCTTTTACTTTGAAGAAATCATGCTATAGA 1759
Db 241 aatggttaatgtgcttctaataaaatggaagtattgtagctggaatgtgatacatgtaaca 300
|||||
Cp 1758 AATGGTTAATGTGCTTCTAATAAATGGAAGTATTGTAGCTGGAATGTGATACATGTAACA 1699
Db 301 gtttaagttcccatgaagggtataaaaatgatgaattgttgtgaagacttagacactggagt 360
|||||
Cp 1698 GTTTAAGTTCCTCATGAAGGTATAAAATGATGAATGTTTGTGAAGACTTAGACACTG-AGT 1640
Db 361 ctcaagtctggagctgatgaagatgttgagataaacagccagctttatcccaacaggggt 420
|||||
Cp 1639 CTCAGTCTGG-AGCTGATGAAGATGTTGAGATAAACAGCCAGCTTTTATCTC-AACAGGGG-T 1583

Db 421 ttgtgaccacaaagttttgg 440
|||||
Cp 1582 TTGTGACCCACAAGTTTGGG 1563

RESULT 6 AA024772 433 bp mRNA EST 13-AUG-1996
LOCUS ze76g10.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
DEFINITION 364962 3'.
ACCESSION AA024772
NID g1489678
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 433)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 289.
Location/Qualifiers
1..433
/organism="Homo sapiens"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCATCTTTT-TTTT-TTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NbHL19W."
/clone="364962"
/clone_lib="Soares fetal heart NbHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
complement(<1..>433)
BASE COUNT 150 a 51 c 87 g 141 t 4 others
ORIGIN

mRNA
Query Match 17.9%; Score 363; DB 156; Length 433;
Best Local Similarity 99.2%; Pred. No. 0.00e+00;
Matches 369; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Db 1 gtaaaacaggatgtaaagtttatatacaagaataataatgtttatctgaaatatttacagt 60
|||||
Cp 2000 GTAAACAGGATGTAAAGTTTATATACAGAATATAATGTTTATCTGAAATATTTACAGT 1941
|||||
Db 61 gttgggttaaagcaaatatttttacaaactttttaaagggtaaactactatgtatatattacaggta 120
|||||
Cp 1940 GTTGGTTAAGCAATATTTTACAACTTTTAAAGGTAAACTACTATGTATATTACAGGTA 1881
|||||

Db 121 agctacaatgggttaatttgcaaaagttaagtaagaaatgtttttaacaaggcttaaag 180
|||||
Cp 1880 AGCTACAATGGGTTTAATTTGCAAAAGTTAAGTAAGAAATGTTTAAACAAGGCTTAAAG 1821
|||||
Db 181 tactcaagtcaattataaaatttatatcttttgccttttacttgaagaatcatgctata 240
|||||
Cp 1820 TACTCAAGTCAATTATAAAATTTATATCTTTTGCCTTTTACTTTGAAGAAATCATGCTATA 1761
|||||
Db 241 gaaatggttaatgtgcttctctaataaaatggaagtattgtagctggaatgtgatacatgtaa 300
|||||
Cp 1760 GAAATGGTTAATGTGCTTCTTAATAAAATGGAAGTATTGTAGTGGATGTGATACATGTAA 1701
|||||
Db 301 cagtttaagttcccatgaagggtataaaatgatgaattgttgaagacttagacacngga 360
|||||
Cp 1700 CAGTTTAAAGTCCCATTTGAAGGTATAAAATGATGAATTGTTGTAAGACTTAGACACTG-A 1642
|||||
Db 361 gtctcagtcngg 372
|||||
Cp 1641 GTCTCAGTCTGG 1630
|||||

RESULT 7
LOCUS H87071 426 bp mRNA EST 21-NOV-1995
DEFINITION ys74d07.r1 Homo sapiens cDNA clone 220525 5' similar to SP:A45054
A45054 FZ-1=PUTATIVE INTERCELLULAR SIGNAL TRANSDUCER OR TRANSMITTER

ACCESSION H87071
NID g1068650
KEYWORDS EST.
SOURCE human clone=220525 primer=M13RP1 library=Soares retina N2b4HR
vector=pT7T3D (Pharmacia) with a modified polylinker host=DH10B
(ampicillin resistant) Rsitel=Not I Rsite2=Eco RI 1st strand cDNA
was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCGCTTTT-TTTT-TTTT-TTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT7T3 vector (Pharmacia). The retinas were
obtained from a 55 year old Caucasian male and total cellular
poly(A)+ RNA was extracted 6 hrs after their removal. The retina
RNA was kindly provided by Roderick R. McInnes M.D. Ph.D. from the
University of Toronto. Library constructed by Bento Soares and
M.Fatima Bonaldo.
Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 426)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 355
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Location/Qualifiers
1..426
/organism="Homo sapiens"
/clone="220525"
<1..>426

FEATURES
source
1..426
/organism="Homo sapiens"
/clone="220525"
<1..>426

BASE COUNT 96 a 128 c 110 g 83 t 9 others

ORIGIN

Query Match 17.7%; Score 359; DB 72; Length 426;
Best Local Similarity 94.8%; Pred. No. 0.00e+00;
Matches 399; Conservative 0; Mismatches 17; Indels

Db 2 acggcatngaataccagaacatcggtgccccacctgctggtggccacgagacatgaagg 61
|||||
QY 407 ACGGCATCGAATACCAGAACATCGGGCTGCCCAACCTGCTGGGCCACGAGACCATGAAG 466
|||||

Db	62	aggtgctggagcaggccgcctggatccccctggtcatgaagcagtgcaccccgcaca	121
Qy	467	AGGTGCTGGAGCAGGCCCGCGCTTGGATCCCGCTGGTCATGAAGCACGTGCACCCGGACA	526

Db	122	ccaagaagtccctgtgctgcgtcttcggcccccgctgcctcgatgacctagacagacca	181
Qy	527	CCAAGAAGTTCCTGTGCTCGCTCCTTCGCCCCCCGGCTGCCTCGATGACCTAGACGAGACCA	586

Dbb 182 tccagccatgccactcgctctg-gtgcaggtgaaaggaccgctgcgcccggtcatgtcc 240
|||||
587 TCCAGCCCATGCCACTCGCTCTGGCTGTCAGGTGAA-GGACCGCTGCCGCCCGGTCAATGTCC 645
|||||

[illegible][illegible]

Db 361 ggatgtgaagctgnaaaattaaaaatgatgatgaacaacgacataaatgggaacgntt 420
|||||
Qy 765 GGTATGTGAAGCCTGCAAAATAAATAATGATGATGA -CAACGACATAATGG-AAACGCTT 822

Db	421	t	421
Qy	823	T	823

```

8
RESULT      H16121      406 bp      mRNA      EST      27-JUN-1995
LOCUS      Y128c05.r1 Homo sapiens cDNA clone 159560 5'.
DEFINITION H16121
ACCESSION  g880941
NID
KEYWORDS   EST.
SOURCE     human clone-159560 library=Soares breast 3NbHBst vector=pT7T3D
           (Pharmacia) with a modified polylinker host=DH10B (ampicillin
           resistant) primer=M13RP1 Rsite1=Not I Rsite2=Eco RI Adult human.
           1st strand cDNA was primed with a Not I - oligo(dT) primer [5',
           TGTTACCAATCTGAAGTGGGAGCGCGCCCTTTTTTTTTTTTTTT 3'],
           double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia),
           digested with Not I and cloned into the Not I and Eco RI sites of a
           modified pT7T3 vector (Pharmacia). Library went through one round
           of normalization to a Cot = 20. Library constructed by Bento Soares
           and M.Fatima Bonaldo.

```

ORGANISM

Homo sapiens

Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 406)

REFERENCE

AUTHORS

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.

TITLE
JOURNAL
The WashU-Merck EST Project
Unpublished (1995)

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
High quality sequence stops: 321
Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES source

```

/organism="Homo sapiens"
/clone="159560"

```

BASE COUNT	130 a	94 c	103 g	77 t	2 others
ORIGIN					

Query Match 16.6%; Score 337; DB 61; Length 406;
Best Local Similarity 98.1%; Pred. No. 0.00e+00;
Matches 368; Conservative 0; Mismatches 2; Indels

Db 1 cceaggacaaacgacaccttgcatccccctcgctagcagcancacacctcctgccagccaccg 60
|||||
692 CCCAGGACAAACGACCTTTGCACTCCCCCTCGCTAGCAGGACCAACCTCCTGCCAGCCACCG 751
Qv

[illegible]

Db	121	tggaacgctttgtaaaaatgatttgcactgaaaaataaaagtgaaggagataaacctaca	180
Qy	812	TGGAACGCTTTGTAAAAATGATTTTGCCTGAAAATAAAGTGAAGGAGATAACCTACA	871

Db 181 tcaacccgagataccacaaatcatcctctggagaccaagagcaagaccatttacaagctgaacg 240
|||||
Ov 872 TCAACCGGAGATACCAAATCATCTCTGGAGACCAAGAGCAAGACCATTTTACAAGCTGAACG 931

Db	241	gtgtgctccgaaggacctgaagaaatcggtgctgtggctcaagaacagcattgcagtgcaca
Qy	932	GTGTGTCGGAAAGGGACCTGAAGAAAATCGGTGCTGTGGCTCAAAGACAGCTTGCAGTGCA

Db 301 cctgtgaggagatgaacgacattcaacgcgcctatctggttcatggggacaggaaacag 360
|||||
Ov 992 CCTGTGAGGAGATGAACGACAT-CAACGCGCCCTATCTGCT-CATGGG-ACAG-AAACAG 104
|||||

Db	361	375
Db	ggttggggagctggt	375
Ov	1048 GGT-GGGGAGCTGGT	106

RESULT	9				
LOCUS	HUM296B07B	340 bp	mrna	EST	29-AUG-1995
DEFINITION	Human aorta cDNA 5'-end GEN-296B07.				
ACCESSION	D62525				
NID	9966299				
KEYWORDS	EST(expressed sequence tag); Human aorta; similar to none(May 29,1995).				
SOURCE	Homo sapiens (library: Clontech human aorta polyA+ mRNA (#6572)) cDNA to mrna.				

ORGANISM	Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;	
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 340)

AUTHORS
Fujiwara, T., Hirano, H., Katagiri, T., Kawai, A., Kuga, Y., Nagata, M.,
Okuno, S., Ozaki, K., Shimizu, F., Shimada, Y., Shinomiya, H.,
Takaichi, A., Takeda, S., Watanabe, T., Takahashi, E.-I., Hirai, Y.,
Maekawa, H., Shin, S. and Nakamura, Y.

TITLE	Unpublished(303)
JOURNAL	Unpublished (1995)
COMMENT	Submitted (30-May-1995) to DDBJ by: Tsutomu Fujiwara

Otsuka GEN Research Institute
Otsuka Pharmaceutical Co., Ltd
463-10 kagasuno Kawauchi-cho
Tokushima, Tokushima
771-01

Japan		Location/Qualifiers	
Phone: 0886-65-2888		1..396	
Fax : 0886-37-1035.		/organism="Mus musculus"	
		/strain="NIH/Swiss"	
		/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. 93 pooled NIH/Swiss 13 day embryo hearts. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTTCTTTT 3'"	
		/clone="536389"	
		/clone_lib="Stratagene mouse heart (#937316)"	
		/sex="pooled"	
		/dev_stage="13 day embryos"	
		/lab_host="SOLR (kanamycin resistant)"	
		<1..>396	
BASE COUNT		93 a	126 c 103 g 74 t
ORIGIN		mRNA	
		ORIGIN	
Query Match		16.2%; Score 328; DB 178; Length 396;	
Best Local Similarity		92.5%; Pred. No. 0.00e+00;	
Matches		368; Conservative 0; Mismatches 28; Indels 2; Gaps 2;	
Db	1	agcaggcgggcgct-ggattccgcgtggtcatgaagcagctgccaccggacacccaagaagt	59
QY	476	AGCAGGCCGGCGCTGGATCCCGTGGTCATGAAGCAGTGCCACCCGGACACCAAGAGT	535
Db	60	tcctgtgctgcgtcttcgcgcctgtctgtctgcgcgccttagatgagaccatccagccgt	119
QY	536	TCCTGTGCTCGCTTCGCCCCCGCTCTGCCTCGATGACCTAGACGACCATCCGCCAT	595
Db	120	gtcactcgctctg-gtgcaggtgaagaccgctgcgcgccttcgcgcgccttcggct	178
QY	596	GCCACTCGCTCGCTGCGTGAAGGACCGCTGCCCGCTCCCGCTCATGTCCGCTTCGGCT	655
Db	179	tcccctggccagacatgctggagtgcgaccgtttccgcgcgcgccttcgcgccttcgcctcc	238
QY	656	TCCCCTGGCCCGACATGCTTGAGTGCAGCCGTTTCCCCCAGGACACGACCTTTGCTCC	715
Db	239	ccctcgctagtagcgaccacctcctcgcgccacagaggaggaagctcccaagggtgtggaag	298
QY	716	CCCTCGCTAGCAGCGACCACTCTCTGCCAGCCACCGAGGAAGTCCAAAGGTATGTGAAG	775
Db	299	cctgcaaaaccaaagaatgaggacgacacacatcatatggaacccctttgtaaaaatgact	358
QY	776	CCTGCAAAATAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	835
Db	359	tcgcactgaaaaatcaaatgaaggagataaactacatc	396
QY	836	TTGCACTGAAAAATAAAAGTGAAGGAGATAACCTACATC	873
RESULT		11	
LOCUS		AA024771 414 bp mRNA EST 13-AUG-1996	
DEFINITION		ze76gl0.r1 Soares fetal heart NbH19W Homo sapiens cDNA clone	
ACCESSION		AA024771	
NID		91489677	
KEYWORDS		EST.	
SOURCE		human.	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS		1 (bases 1 to 414)	
TITLE		The WashU-Merck EST Project	
JOURNAL		Unpublished (1995)	
COMMENT		Contact: Wilson RK	

Japan		Location/Qualifiers	
Phone: 0886-65-2888		1..340	
Fax : 0886-37-1035.		/organism="Homo sapiens"	
		/clone_lib="Clontech human aorta polyA+ mRNA (#6572)"	
BASE COUNT		113 a	64 c 38 g 122 t
ORIGIN		16.3%; Score 331; DB 48; Length 340;	
Best Local Similarity		98.5%; Pred. No. 0.00e+00;	
Matches		335; Conservative 3; Mismatches 1; Indels 1; Gaps 1;	
Db	1	catcttcatcagctccagactgagactcaggtgtctaagtcttacacaattcatctttt	60
QY	1616	CATCTTCATCAGCTCCAGACTGAGACTCAGTGTCTAAGTCTTACACAATTCATCATTTT	1675
Db	61	ataccttcaatgggaacttaaaactgttacatgtatcacattccagctacaatacttccat	120
QY	1676	ATACCTTCAATGGGAACCTTAAACTGTATACATGTATCACATTCACGTTACAATACTTCCAT	1735
Db	121	ttatttagaggcacatttaaccatttctatagcatgatttcttccaagtaaaaggcaaaagat	180
QY	1736	TTATTAGAAGCACATTAAACCATTTCTATAGCATGATTTCTTCAAGTAAAGGCAAAAGAT	1795
Db	181	ataaatcttataaattgacttgacttgaacttgaagcctgtgttaaaacatttcttacttaact	240
QY	1796	ATAAATTTTATAATTGACTTGAGTACTTTAAGCCTTGTTTAAACATTTCTTACTTAACT	1855
Db	241	tttggcaaatataaccattgtagcttacctgtaataatatacatagtrgtttacctttaaaa	300
QY	1856	TTTG-CAAAATAACCCATTGTAGCTTACCTGTAATATACATAGTAGTTACCTTTAAAA	1914
Db	301	gttgtaaaaartattgcttttraccacacactgtaaatatttc	340
QY	1915	GTTGTAAAAATATTGCTTTAACCAACACTGTAATATTTC	1954
RESULT		10	
LOCUS		AA073862 396 bp mRNA EST 04-OCT-1996	
DEFINITION		mm97f07.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone	
ACCESSION		AA073862	
NID		91595602	
KEYWORDS		EST.	
SOURCE		house mouse.	
ORGANISM		Mus musculus	
REFERENCE		Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS		1 (bases 1 to 396)	
TITLE		The WashU-HMI Mouse EST Project	
JOURNAL		Unpublished (1996)	
COMMENT		Contact: Marra M/Mouse EST Project	
		WashU-HMI Mouse EST Project	
		Washington University School of Medicine	
		4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108	
		Tel: 314 286 1800	
		Fax: 314 286 1810	
		Email: mouseest@watson.wustl.edu	
		This clone is available royalty-free through LLNL; contact the	
		IMAGE Consortium (info@image.llnl.gov) for further information.	
		MGI:323325	
		Seq primer: primer name ambiguous	
		High quality sequence stop: 390.	

Db 422 ttaaaagtggttaa 434
|||||
QY 1909 TTAAGAAGTTGTAA 1921

RESULT 13
LOCUS H15818 469 bp mRNA EST 27-JUN-1995
DEFINITION Y128C05.s1 Homo sapiens cDNA clone 159560 3'.
ACCESSION H15818
NID g880638
KEYWORDS EST.
SOURCE human clone=159560 library=Soares breast 3NbHBst vector=pf7T3D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) primer=Promega -2lml3 Rsite1-Not I Rsite2=Eco RI Adult human. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTACCAATCTGAAGTGGAGCGGCCCTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pf7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 20. Library constructed by Bento Soares and M.Fatima Bonaldo.

ORGANISM Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 469)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)

REFERENCE
AUTHORS Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 340
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES
source
1..469
/organism="Homo sapiens"
/clone="159560"

BASE COUNT 144 a 65 c 106 g 150 t 4 others
ORIGIN

Query Match 15.8%; Score 321; DB 61; Length 469;
Best Local Similarity 97.1%; Pred. No. 0.00e+00;
Matches 370; Conservative 0; Mismatches 3; Indels 8; Gaps 8;

Db 1 aggtataaacaggatttaaaagtttatatacaagaataatgtttatctgaaatatata 60
|||||
Cp 2002 AGGTAAACAGGATGTAAAGTTTATATACAGAATATAATGTTATCTGAAATATTACA 1943
|||||

Db 61 gtgttggttaaagcaatatatttttacaacttttaaaggtaaactactatgtatattacagg 120
|||||

Cp 1942 GTGTTGGTTAAGCAATATTTTACAACCTTTTAAAGGTAACACTACTATGTATATTACAGG 1883
|||||

Db 121 taagctacaatg9gggtttaatttgcaaaaagttaagtaagaaatggttttaacaaggcttaa 180
|||||

Cp 1882 TAAGCTACAATGGGTTTAATTGCAAAAAGTTAAGTAAGAAATGTTTAAACAAGGCTTAA 1823
|||||

Db 181 agtactcaagtcaattataaaatttatatcttttgccttttacttgaagaaatcatgcta 240
|||||

Cp 1822 AGTACTCAAGTCAATTATAAAATTTATATCTTTTGCCTTTTACTTGAAGAAATCATGCTA 1763
|||||

Db 241 tagaaatggttaatgtgcttctaataaaatgggaagtattgttaggctgggaatgtgataca 300
|||||
Cp 1762 TAGAAATGGTTAATGTGCTTCTAATAAATGG-AAGTATTGTAG-CTGG-AATGTGATACA 1706
|||||

Db 301 tgtaacaggtttaaaggtttcccatgaagggtataaaatgatgaattgttgaaggncttag 360
|||||
Cp 1705 TGTAACAGTTTAAG-TTCCCATTTGAAGG-TATAAAATGATGAATTGTTGTAG-ACTTAG 1649
|||||

Db 361 gacactgaggtctcagtcngg 381
|||||
Cp 1648 -ACACTGAG-TCTCAGTCTGG 1630
|||||

RESULT 14
LOCUS N32424 326 bp mRNA EST 10-JAN-1996
DEFINITION YW83e07.s1 Homo sapiens cDNA clone 258852 3'.
ACCESSION N32424
NID g1152823
KEYWORDS EST.
SOURCE human clone=258852 primer=ml3 -40 forward library=Soares placenta 8to9weeks 2NbHP8to9W vector=pf7T3D (Pharmacia) with a modified polylinker host=DH10B (ampicillin resistant) Rsite1=Not I Rsite2=Eco RI two placentae: one from 8 weeks and another from 9 weeks post conception. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGGAGCGCGCGGATTTTTTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pf7T3 vector (Pharmacia). Library constructed by Bento Soares and M. Fatima Bonaldo.

ORGANISM Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 326)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)

REFERENCE
AUTHORS Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 312
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES
source
Location/Qualifiers
1..326
/organism="Homo sapiens"
/clone="258852"
<1..>326

BASE COUNT 123 a 33 c 54 g 116 t
ORIGIN

Query Match 15.7%; Score 319; DB 92; Length 326;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 325; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 1 taaaacaggatgataaagtttatatacaagaataataatgtttatctgaaatatattacagt 60
|||||
Cp 1999 TAAAACAGGATGTAAAGTTTATATACAGAATATAATGTTTATCTGAAATATTACAGTG 1940
|||||

Db 61 ttggttaaagcaatatattttacaacttttaaaaggttaaactactatgtatatattacaggtaa 120
|||||

Cp 1939 TTGGTTAAAGCAATATTTTACAACCTTTTAAAGGTAAACTACTATGTATATTACAGGTAA 1880
Db 121 gctacaatgggtttaattgcaaaagttaagtaagaatggttttaaacacaggccttaaagt 180
Cp 1879 GCTACAATGGGTTTAATTGCAAAAGTTAACTAAGAATGTTTTTAAACAACGGCTTAAAGT 1820
Db 181 actcaagtcaattatataaaatttatatcttttgcctttttacttgaaagaatacatgctatac 240
Cp 1819 ACTCAAGTCAATTATATAAATTTATATCTTTTGCCCTTTTACTTGAAGAAATCATGCTATA- 1761
Db 241 gaaatgggttaatgtgcttctctaataaataaggaagtattgttagctggaaatgtgatacatgtaa 300
Cp 1760 GAAATGGTTAATGTGCTTCTTAATAAATGGAAGTATTGTAGCTGGAATGTGATACATGTAA 1701
Db 301 cagtttaagttcccatggaaggata 326
Cp 1700 CAGTTAAGTTCCTCCATTGAAGGTATA 1675

RESULT 15
LOCUS HUM230F10B 330 bp mRNA EST 28-AUG-1995
DEFINITION Human aorta cDNA 5'-end GEN-230F10.
ACCESSION D62000
NID g965776
KEYWORDS EST(expressed sequence tag); Human aorta; similar to none(May 29,1995).
SOURCE Homo sapiens (library: Clontech human aorta polyA+ mRNA (#6572))
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 330)
AUTHORS Fujiwara,T., Hirano,H., Katagiri,T., Kawai,A., Kuga,Y., Nagata,M., Okuno,S., Ozaki,K., Shimizu,F., Shimada,Y., Shinomiya,H., Takaichi,A., Takeda,S., Watanabe,T., Takahashi,E.-I., Hirai,Y., Maekawa,H., Shin,S. and Nakamura,Y.
TITLE Unpublished(303)
JOURNAL Unpublished (1995)
COMMENT Submitted (30-May-1995) to DDBJ by: Tsutomu Fujiwara
Otsuka GEN Research Institute
Otsuka Pharmaceutical Co.,Ltd
463-10 kagasuno Kawauchi-cho
Tokushima, Tokushima 771-01
Japan
Phone: 0886-65-2888
Fax : 0886-37-1035.

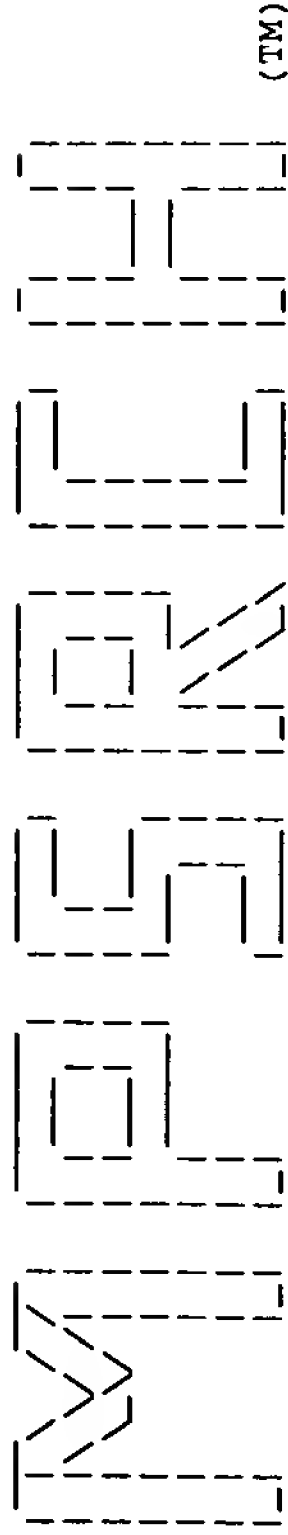
FEATURES
source 1..330
/organism="Homo sapiens"
/clone_lib="Clontech human aorta polyA+ mRNA (#6572)"
BASE COUNT 116 a 57 c 32 g 118 t 7 others
ORIGIN

Query Match 15.6%; Score 316; DB 48; Length 330;
Best Local Similarity 97.8%; Pred. No. 0.00e+00;
Matches 316; Conservative 6; Mismatches 0; Indels 1; Gaps 1;

Db 9 caacaattcatcattttataaccttcaatgggaacttaaaactgttacatgtatcacattcc 68
QY 1659 CAACAATTTCATCATTTTATACCTTCAATGGGAACCTTAAACTGTTACATGTATCACATTCC 1718
Db 69 agctacaatacttccattttatttagaagcacattaaaccatttctctatagcatgtttcttca 128
QY 1719 AGCTACAATACTTCCATTTATTAGAGCACATTAAACCATTTCTATAGCATGATTCTTCA 1778
Db 129 agtaaaaggcaaaagatataaaatyvata-ttsacttgagtacttttaagccttgtttaaa 187
QY 1779 AGTAAAGGCAAAAGATATAAATTTTATAATTGACTTGAGTACTTTAAGCCTTGTTTAAA 1838
Db 198 acatttcttacttaacttttgcaaaattaaacccattgttagcttacctgtaatatacatag 247

QY 1839 ACATTTCTTACTTAACTTTTGCAAATTAAACCCATTGTAGCTTACCTGTAAATATACATAG 1898
Db 248 tagtttacccctttaavagttgtaaaaaatattgcttttaaccaaacactgttaaatatttcagat 307
QY 1899 TAGTTTACCTTTTAAAGTTGTAAAAATATTGCTTTTAAACCAACACTGTAAATATTTCAGAT 1958
Db 308 aaacattatatattcyygtatataa 330
QY 1959 AAACATTATATTCTTGTATATAA 1981

Search completed: Tue Dec 9 09:09:41 1997
Job time : 1565 secs.



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPSrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 9 09:10:03 1997; MasPar time 537.81 Seconds
Tabular output not generated. 1035.748 Million cell updates/sec

Title: >US-08-848-439-1
Description: (1-2027) from US08848439.seq
Perfect Score: 2027
N.A. Sequence: 1 GAATTCGGCCTTCATGGCCT.....AAAAAAGCGGCCGC 2027
Comp: CTTAAGCCGGAAGTACCGGA.....TTTTTTTTCGCCGGCG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 359085 seqs, 137405154 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: EST-STS-THREE
1:EST199 2:EST200 3:EST201 4:EST202 5:EST203 6:EST204
7:EST205 8:EST206 9:EST207 10:EST208 11:EST209 12:EST210
13:EST211 14:EST212 15:EST213 16:EST214 17:EST215
18:EST216 19:EST217 20:EST218 21:EST219 22:EST220
23:EST221 24:EST222 25:EST223 26:EST224 27:EST225
28:EST226 29:EST227 30:EST228 31:EST229 32:EST230
33:EST231 34:EST232 35:EST233 36:STS1 37:STS2 38:STS3
39:STS4 40:STS5 41:STS6 42:STS7 43:STS8 44:STS9 45:STS10
46:STS11 47:STS12 48:STS13
EST-STS-FOUR
49:gnEST1 50:gnEST2 51:gnEST3 52:gnEST4 53:gnEST5
54:gnEST6 55:gnEST7 56:gnEST8 57:gnEST9 58:gnEST10
59:gnEST11 60:gnEST12 61:gnEST13 62:gnEST14 63:gnEST15
64:gnEST16 65:gnEST17 66:gnEST18 67:gnEST19 68:gnEST20
69:gnEST21 70:gnEST22 71:gnEST23 72:gnEST24 73:gnEST25
74:gnSTS 75:enEST1 76:enEST2 77:enEST3 78:enEST4
79:enEST5 80:enEST6 81:enEST7 82:enEST8 83:enEST9
84:enEST10 85:enEST11 86:enEST12 87:enEST13 88:enEST14
89:enEST15 90:enEST16 91:enEST17 92:enEST18 93:enEST19
94:enEST20 95:enEST21 96:enEST22 97:enEST23 98:enEST24
99:enEST25 100:enEST26 101:enEST27 102:enEST28
103:enEST29 104:enEST30 105:enEST31 106:enSTS 107:ueEST1
108:ueEST2

Database:

ALIGNMENTS

RESULT 1
LOCUS AA071469 505 bp mRNA EST 01-FEB-1997
DEFINITION ze88f06.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone
366083 3'
ACCESSION AA071469
NID g1578900
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 505)
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfling, T., Tan, F., Trevaskis, E.,
Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.
TITLE WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT

Statistics: Mean 12.177; Variance 2.973; scale 4.096

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine

No.	Score	Match	Length	DB	ID	Description	Pred. No.
C 1	479	23.6	505	32	AA071469	ze88f06.s1 Soares fet	0.00e+00
2	472	23.3	505	2	W92531	ze03g10.r1 Soares fet	0.00e+00
3	472	23.3	505	90	HSW5314	ze03g10.r1 Soares fet	0.00e+00
4	366	18.1	458	33	AA105749	ml84a03.r1 Stratagene	0.00e+00
5	328	16.2	396	54	AA073862	mm97f07.r1 Stratagene	0.00e+00
C 6	256	12.6	262	11	AA176165	zp23h12.s1 Stratagene	0.00e+00
7	255	12.6	287	32	AA082155	ze88f06.r1 Soares fet	0.00e+00
8	255	12.6	394	101	MMAA60087	va52g08.r1 Soares mou	0.00e+00
9	242	11.9	523	56	AA122822	mr03c11.r1 Soares mou	0.00e+00
10	242	11.9	523	6	AA122822	mr03c11.r1 Soares mou	0.00e+00
11	209	10.3	299	94	MM3451	mb41f02.r1 Soares mou	7.57e-286
12	156	7.7	184	20	AA209833	mo90h01.r1 Beddington	8.95e-198
13	114	5.6	145	64	AA230899	mw14c07.r1 Soares mou	1.03e-129
14	114	5.6	145	99	MMAA30899	mw14c07.r1 Soares mou	1.03e-129
15	86	4.2	120	20	AA206685	zg51e09.r1 Stratagene	7.12e-86
C 16	47	2.3	502	21	AA020088	mh49c08.r1 Soares mou	2.88e-29
17	32	1.6	496	10	AA013810	mh06e01.r1 Soares mou	1.13e-10
C 18	31	1.5	432	29	AA030216	mh87g08.r1 Soares mou	1.47e-09
C 19	29	1.4	156	36	DM367S	D. melanogaster STS d	2.13e-07
20	28	1.4	262	11	AA176165	zp23h12.s1 Stratagene	2.36e-06
C 21	28	1.4	292	81	HS1145682	nc19a09.r1 NCI CGAP P	2.36e-06
C 22	28	1.4	292	63	AA226250	nc19a09.r1 NCI CGAP P	2.36e-06
23	28	1.4	376	9	AA133408	zn92f01.r1 Stratagene	2.36e-06
C 24	28	1.4	397	40	G14542	human STS SHGC-11394.	2.36e-06
25	28	1.4	400	40	G13598	human STS SHGC-11851.	2.36e-06
C 26	29	1.4	415	14	AA187208	zp59c07.s1 Stratagene	2.13e-07
27	28	1.4	505	32	AA071469	ze88f06.s1 Soares fet	2.36e-06
C 28	26	1.3	163	87	HS439337	zb96f03.s1 Soares par	2.40e-04
29	26	1.3	205	1	N97804	1287C3 czappFDD2.1, D	2.40e-04
30	26	1.3	449	100	MMAA41089	mu32h11.r1 Soares 2Nb	2.40e-04
31	26	1.3	449	62	AA204488	mu32h11.r1 Soares 2Nb	2.40e-04
32	26	1.3	449	19	AA204488	mu32h11.r1 Soares 2Nb	2.40e-04
C 33	27	1.3	482	21	AA019433	ze55g01.s1 Soares ret	2.46e-05
C 34	26	1.3	543	32	AA089167	ml85b01.r1 Stratagene	2.40e-04
C 35	26	1.3	570	101	MMAA59657	va36a10.r1 Soares mou	2.40e-04
36	25	1.2	158	36	DM48A3T	D. melanogaster STS d	2.19e-03
37	25	1.2	305	21	AA019656	ze61f11.s1 Soares ret	2.19e-03
C 38	25	1.2	327	1	N97836	2124C3 czappFDD2.1, D	2.19e-03
C 39	25	1.2	330	1	N97581	1321C3 czappFDD2.1, D	2.19e-03
40	25	1.2	366	67	AA238341	mx95c11.r1 Soares mou	2.19e-03
41	25	1.2	375	24	AA216813	mu86h01.r1 Soares mou	2.19e-03
C 42	25	1.2	379	3	AA084261	zn03e11.s1 Stratagene	2.19e-03
43	25	1.2	413	95	MM4855	ma53e06.r1 Soares mou	2.19e-03
C 44	25	1.2	544	13	AA186000	mu56h10.r1 Soares mou	2.19e-03
C 45	25	1.2	1171	95	MM52713	mc03e12.r1 Soares mou	2.19e-03


```

RESULT 6
LOCUS AAL176165 262 bp mRNA EST 30-DEC-1996
DEFINITION zp23hl2.s1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA
Clone 610343 3'.
ACCESSION AAL176165
NID 91757287
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 262)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
WashU-Merck EST Project
WashU-Merck EST Project
Unpublished (1995)
TITLE
JOURNAL
COMMENT
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 161.
Location/Qualifiers
FEATURES
source
1..262
/organism="Homo sapiens"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT. NT2
cells (Ntera-2/cl.D1) induced with Retinoic Acid for 24
hours. Average insert size: 1.5 kb; Uni-ZAP XR Vector; -5'
adaptor sequence: 5' GAATTCGGCAGGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTT TTTT TTTT TTTT TTTT 3'"
/clone="610343"
/clone_lib="Stratagene neuroepithelium (#937231)"
/dev_stage="Ntera-2/RA neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
complement(<1..>262)
mRNA 100 a 27 c 40 g 91 t 4 others
BASE COUNT
ORIGIN
Query Match 12.6%; Score 256; DB 11; Length 262;
Best Local Similarity 98.1%; Pred. No. 0.00e+00;
Matches 257; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Db 1 ggtaaaacaggatgtaaagtttatatacaagaataataatgtttatctgaaatattacag 60
|||||
Cp 2001 GGTAAACACAGGATGTAAAGTTTATATACAAGAATATAATGTTTATCTGAAATATTACAG 1942
Db 61 tgttggttaaagcaatatattttacaacttttaaaaggtaaactactatgtatatattacaggt 120
|||||
Cp 1941 TGTGGTTAAAGCAATATTTTACAACTTTTAAAGTTAACTACTATGTATATTACAGGT 1882
Db 121 aagctacaaatgggtttaatttgcaaaagttaagtaagaaatgttttaacaagaagcctaaa 180
|||||
Cp 1881 AAGCTACAATGGGTTTAATTGCAAAAGTTAAGTAAGAAATGTTTAAACAAGGCTTAAA 1822
Db 181 gtactcaagtcaattataaaaatttatancctttngccttttacttgaaagaatcatgctat 240
|||||
Cp 1821 GTACTCAAGTCAATTATAAAATTTATATCTTTTGCCTTTTACTTGAAGAAATCATGCTAT 1762
Db 241 agaaatggntaatgtncctcta 262
|||||
Cp 1761 AGAAATGGTTAATGTGCTTCTA 1740
RESULT 7
```

```

LOCUS AA082155 287 bp mRNA EST 01-FEB-1997
DEFINITION ze88f06.r1 Soares fetal heart NbHL19W Homo sapiens cDNA clone
366083 5'.
ACCESSION AA082155
NID 91624408
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 287)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
WashU-Merck EST Project
WashU-Merck EST Project
Unpublished (1995)
TITLE
JOURNAL
COMMENT
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1171 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 262.
Location/Qualifiers
FEATURES
source
1..287
/organism="Homo sapiens"
/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGGAGCGCGCCGACATCTTTT TTTT TTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT7T3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NbHL19W."
/clone="366083"
/clone_lib="Soares fetal heart NbHL19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
<1..>287
mRNA 62 a 84 c 85 g 51 t 5 others
BASE COUNT
ORIGIN
Query Match 12.6%; Score 255; DB 32; Length 287;
Best Local Similarity 97.5%; Pred. No. 0.00e+00;
Matches 273; Conservative 0; Mismatches 4; Indels 3; Gaps 3;
Db 1 ggctcaaaagacagcttgacgtgcacctgtgaggagatgaacacatcaacgcgcctatc 60
|||||
QY 968 GGCTCAAAGACAGCTTGCAGTGCACCTGTGAGGAGATGAACGACATCAACGGCCCTATC 1027
Db 61 tggctatgggacagaacacaggggtggggagctggtgatcacctcggtgaagcggtagcaga 120
|||||
QY 1028 TGGTCATGGGACAGAAACAGGGTGGGAGCTGGTGATCACCTCGGTGAAGCGGTGGCAGA 1087
Db 121 aggggcagagagagttcaagcgcatctncccggnngnatccgcaagctgcagtgctagtc 180
|||||
QY 1088 AGGGGCAGAGAGAGTTCAAGCGCATCTCCCCG-CAGCATCCGCAAGCTGCAGTGTAGTCC 1146
Db 181 cggcatcctgatggtcccgacag-cctgctcccagagcacggctgaccatttctgctccgg 239
|||||
QY 1147 CGGCATCCTGATGCTCCGACAGGCCCTGCTCCAGAGCACGGCTGACCATTTCGCTCCGG 1206
```


Db 240 gatctcagctcccggttccc-aagcacactcctagtgctc 278
|||||
QY 1207 GATCTCAGCTCCCGTTCCCAAGCACACTCCTAGCTGCTC 1246

RESULT 8
ID MAA60087 standard; RNA; EST; 394 BP.
AC AA260087;
NI g1896590
DT 19-MAR-1997 (Rel. 51, Created)
DT 19-MAR-1997 (Rel. 51, Last updated, Version 1)
DE va52g08.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone 735038 5'.

KW EST.
OS Mus musculus (house mouse)
OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
OC Mus.
RN [1]
RP 1-394
RA Marra M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA Waterston R.;
RT "The WashU-HHMI Mouse EST Project";
RL Unpublished.

CC Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC Washington University School of MedicineP 4444 Forest Park Parkway,
CC Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC Email: mouseest@watson.wustl.edu This clone is available
CC royalty-free through LLNL ; contact the IMAGE Consortium
CC (info@image.llnl.gov) for further information. MGI:452086 Seq
CC primer: -28ml3 rev2 ET from Amersham.
FH Key Location/Qualifiers

FH source 1..394
FT /organism="Mus musculus"
FT /strain="C57BL/6J"
FT /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
FT polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
FT was primed with a Not I - oligo(dT) primer [5'
FT TGTTACCAATCTGAAGTGGAGCGGCCCTTATTTTTTT 3'], o
n
FT total mouse RNA [providedby Minoru Ko, Wayne State Univ.];
FT double-stranded cDNA was ligated to Eco RI adaptors
FT (Pharmacia), digested with Not I and cloned into the Not I
FT and Eco RI sites of the modified pT7T3 vector. Library wen
t
FT through one round of normalization, and was constructed by
FT Bento Soares and M. Fatima Bonaldo."
FT /clone="735038"
FT /clone_lib="Soares mouse 3NME12 5"
FT /sex="unknown"
FT /dev_stage="12.5dpc total fetus"
FT /lab_host="DH10B"
FT mRNA <1..>394
SQ Sequence 394 BP; 104 A; 96 C; 118 G; 76 T; 0 other;

Query Match 12.6%; Score 255; DB 101; Length 394;
Best Local Similarity 94.1%; Pred. No. 0.00e+00;
Matches 272; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Db 1 aaggagataacgtacatcaacagagacacccaagatcatcctggagacaaagagcacc 60
|||||
QY 856 AAGGAGATAACCTACATCAACCGAGATACCAAAATCATCTGGAGACCAAGACACC 915
Db 61 attacaagctgaacggcgtgtccgaaggacacctgaagaaatccgtgctggtcctcaaa 120
|||||
QY 916 ATTTACAAGCTGAACGGTGTGTCCGAAGGACCTGAAGAAATCGGTGCTGCTGCTCAA 975
Db 121 gacagcctgcagtgccacctgtgaggagatgaacgacatcaacgctccgtatctggtcatg 180
", ,
|||||

QY 976 GACAGCTTGCACTGTGAGGAGATGAACGACATCAACGGCCCTATCTGGTCATG 1035
Db 181 ggacagaagcagggcgagctggtgatcacctccgtgaacaggtgcagaagggccag 240
|||||
QY 1036 GGACAGAAACAGGGTGGGAGCTGGTGATCACCTCGTGAAGCGGTGGCAGAAGGGCAG 1095
Db 241 agagagttcaagcgcatctcccgagcatccgcaagctgcaatgtagt 289
|||||
QY 1096 AGAGAGTTCAAGCGCATCTCCCGCAGCATCCGCAAGCTGCAGTGCTAGT 1144

RESULT 9
LOCUS AA122822 523 bp mRNA EST 16-FEB-1997
DEFINITION mr03cl1.r1 Soares mouse 3NbMS Mus musculus cDNA clone 596372 5'.
ACCESSION AA122822
NID g1681779
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.

REFERENCE 1 (bases 1 to 523)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HHMI Mouse EST Project
Unpublished (1996)

TITLE
JOURNAL
COMMENT
Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington Unlversity School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:361804

Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 469.
FEATURES
source
1..523
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGGCCGTGTTT 3']
3'; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. RNA
provided by Dr. Bertrand Jordan. Library went through
three rounds of normalization, and was constructed by
Bento Soares and M.Fatima Bonaldo."
/clone="596372"
/clone_lib="Soares mouse 3NbMS"
/sex="male"
/tissue_type="Spleen"
/dev_stage="4 weeks"
/lab_host="DH10B"
mRNA <1..>523
BASE COUNT 126 a 142 c 129 g 126 t
ORIGIN

Query Match 11.9%; Score 242; DB 56; Length 523;
Best Local Similarity 80.2%; Pred. No. 0.00e+00;
Matches 408; Conservative 0; Mismatches 88; Indels 13; Gaps 5;
Db 1 tccggtgctgtggctcaaaagacagcctgcagtcgacctgtgaggagatgaacgacatcaac 60
|||

REFERENCE 1 (bases 1 to 145)
AUTHORS Marra M., Hillier L., Allen M., Bowles M., Dietrich N., Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J., Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K., Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R. and Waterston R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:410372
Putative full length read
vector to vector length is 146
Seq primer: -28m13 rev2 ET from Amersham.
Location/Qualifiers
1..145
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCCTTATTTT TTTT 3'], on total mouse RNA [provided by Minoru Ko, Wayne State Univ.]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
/clone="670668"
/clone_lib="Soares mouse 3NME12 5"
/sex="unknown"
/dev_stage="12.5dpc total fetus"
/lab_host="DH10B"
<1..>145
mRNA
BASE COUNT 51 a 22 c 19 g 53 t
ORIGIN
Query Match 5.6%; Score 114; DB 64; Length 145;
Best Local Similarity 92.5%; Pred. No. 1.03e-129;
Matches 135; Conservative 0; Mismatches 9; Indels 2; Gaps 2;
Db 1 gcaaaactaaatc-attgtagcttacgtgtaataatcgtagtagttgacctggaaaagtgtg 59
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1859 GCAAAATTAACCCATTGTAGCTTACCTGTAATATACATAGTAGTTTACCTTTAAAAGTTG 1918
Db 60 taaaaatattgctttaaccgacactgtaataatttcagataaacaattattcttctgtat 119
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1919 TAAAAATATTGCTTTAACCACACTGTAAATATTTCAGATAAACATTATATTCTT-GTAT 1977
Db 120 ataaactttacatcttctgtgtaccta 145
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1978 ATAAACTTTACATCCTGTTTACCTA 2003
RESULT 14
ID MAA30899 standard; RNA; EST; 145 BP.
AC AA230899;
NI g1853194
DT 27-FEB-1997 (Rel. 51, Created)
DT 27-FEB-1997 (Rel. 51, Last updated, Version 1)
DE mw14c07.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone 670668 5'.
KW EST.
OS Mus musculus (house mouse)
OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

OC Mus.
RN [1]
RP 1-145
RA Marra M., Hillier L., Allen M., Bowles M., Dietrich N., Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J., Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K., Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R., Waterston R.,
RT "The WashU-HHMI Mouse EST Project";
RL Unpublished.
CC Contact: Marra M/Mouse EST Project WashU-HHMI Mouse EST Project
CC Washington University School of Medicine 4444 Forest Park Parkway,
CC Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC Email: mouseest@watson.wustl.edu This clone is available
CC royalty-free through LLNL; contact the IMAGE Consortium
CC (info@image.llnl.gov) for further information. MGI:410372 Putative
CC full length read vector to vector length is 146 Seq primer: -28m13
CC rev2 ET from Amersham.
FH Key Location/Qualifiers
FT 1..145
FT source
FT /organism="Mus musculus"
FT /strain="C57BL/6J"
FT /note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTGGAGCGCGCCTTATTTT TTTT 3'], on total mouse RNA [provided by Minoru Ko, Wayne State Univ.]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."
FT /clone="670668"
FT /clone_lib="Soares mouse 3NME12 5"
FT /sex="unknown"
FT /dev_stage="12.5dpc total fetus"
FT /lab_host="DH10B"
FT <1..>145
SQ Sequence 145 BP; 51 A; 22 C; 19 G; 53 T; 0 other;
Query Match 5.6%; Score 114; DB 99; Length 145;
Best Local Similarity 92.5%; Pred. No. 1.03e-129;
Matches 135; Conservative 0; Mismatches 9; Indels 2; Gaps 2;
Db 1 gcaaaactaaatc-attgtagcttacgtgtaataatcgtagtagttgacctggaaaagtgtg 59
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1859 GCAAAATTAACCCATTGTAGCTTACCTGTAATATACATAGTAGTTTACCTTTAAAAGTTG 1918
Db 60 taaaaatattgctttaaccgacactgtaataatttcagataaacaattattcttctgtat 119
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1919 TAAAAATATTGCTTTAACCACACTGTAAATATTTCAGATAAACATTATATTCTT-GTAT 1977
Db 120 ataaactttacatcttctgtgtaccta 145
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1978 ATAAACTTTACATCCTGTTTACCTA 2003
RESULT 15
LOCUS AA206685 120 bp mRNA EST 27-JAN-1997
DEFINITION zq51e09.r1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA clone 645160 5'.
ACCESSION AA206685
NID g1802055
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 120)
AUTHORS Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,

Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
WashU-Merck EST Project
Unpublished (1995)

TITLE
JOURNAL
COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 67.

FEATURES
Location/Qualifiers

1..120
/organism="Homo sapiens"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT. NT2
cells (Ntera-2/cl.D1) induced with Retinoic Acid for 24
hours. Average insert size: 1.5 kb; Uni-ZAP XR Vector; -5'
adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
/clone="645160"
/clone_lib="Stratagene neuroepithelium (#937231)"
/dev_stage="Ntera-2/RA neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
<1..>120

BASE COUNT 12 a 54 C 35 g 18 t 1 others
ORIGIN

Query Match 4.2%; Score 86; DB 20; Length 120;
Best Local Similarity 89.3%; Pred. No. 7.12e-86;
Matches 108; Conservative 0; Mismatches 11; Indels 2; Gaps 2;

Db 2 ccccccgagctgcg-gcgggcttgagcncctgccccgcctgtcctcccggtgtcccgct 60
|||||
Qy 48 CCCCCCGGAGCTGGCGCGGGGCTTGACAGGCCCTCGCCCGCGCTCCTCCCGGTGTCGGCT 107

Db 61 tctccgcgccccagccgcccgg-tgccagcttttggggcccccgagtcgcaaccaacgaag 119
|||||
Qy 108 TCTCCGCGCCCCAGCGCGCGGCTGCCAGCTTTTCGGGGCCCCCGAGTCGCACCCAGCGAAG 167

Db 120 a 120
Qy 168 A 168

Search completed: Tue Dec 9 09:22:41 1997
Job time : 758 secs.

1899	5AGTTGACCTGGAAAAAGTTGTAAAAATATTGCTTTAACCGACACTGTAAATATTTCAGAT	1747
1899	TAGTTTACCTTTAAAGAGTTGTAAAAATATTGCTTTAACCAACACTGTAAATATTTCAGAT	1958
1748	aaacattatattcttctgtgtatatataaaacttt	1776
1959	AAACATTATATCTTT-GTATATAAACTTT	1986

RESULT	2	
LOCUS	HSU43318	2334 bp mRNA PRI 24-FEB-1996
DEFINITION	Human putative transmembrane receptor (frizzled 5) mRNA, complete cds.	
ACCESSION	U43318	
NID	g1151251	
KEYWORDS		
SOURCE	human.	
ORGANISM	Homo sapiens	
	Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 2334)	
AUTHORS	Wang, Y., Macke, J.P., Abella, B.S., Andreasson, K., Worley, P., Gilbert, D.J., Copeland, N.G., Jenkins, N.A. and Nathans, J.	
TITLE	A large family of putative transmembrane receptors homologous to the product of the Drosophila tissue polarity gene frizzled	
JOURNAL	J. Biol. Chem.	271 (8), 4468-4476 (1996)
MEDLINE	96224032	
REFERENCE	2 (bases 1 to 2334)	
AUTHORS	Abella, B., Wang, Y., Macke, J.P. and Nathans, J.	
TITLE	Direct Submission	
JOURNAL	Submitted (14-DEC-1995) Jeremy Nathans, Molecular Biology and Genetics, Johns Hopkins Medical School, 725 N. Wolfe Street, Baltimore, MD 21205	

FEATURES	source	CDS	Location/Qualifiers
			1..2334
			/organism="Homo sapiens"
			/chromosome="2"
			/map="2q33-34"
			/tissue_type="retina"
			321..2078
			/gene="frizzled 5"
			/note="putative transmembrane receptor"
			/codon_start=1
			/product="transmembrane receptor"
			/db_xref="PID:g1151252"
			/translation="MARPDSPAPSLLLLLLQAQLVGRAAAASKAPVCQEITVPMCRGI GYNLTMPNQFNHDTQDEAGLEQHFPLVEIQSPDLRFFLCMTPTICLPDYHKPL PPCRSVCERAKAGCSPLMRQYGFAPWERMSCDRLPVLGRDAEVLCDYNRSEATTAPP RPFPAKPTLPDPPGAPASGEGCPAGGPFVCKREPFVILKESHPLYNKVRTGQVPNC AVCYQPSFSADERTFATFWIGLSVLCFISTATVFLIDMTFRYPERPFIIFLSA CPLCVSLGFLVRLVVGHASVACSRHNHHIHYETTGPALECTIVFLVLYFFGMASSIIWY ILSLTWFLAAAMKWGNEIAGYGYFHAAWLIPSVKSIATALALSSVDGDPVAGICYV GNQNLNSLRREVLPGLVLYLLVGTLLFLLAGFVSLFRIRSVIKQGGTKDKLEKLMIRI GIFTLLYTVPASIVVACYLEQHYRESWEAALTACPCGHDGTGQPRAKPEYFWVLMKLYF MCLVVGITSGVMIWSGKTVESWRRFTSRCCCRPRRGHKSGGAMAAGDYPEASAALTGR TGPPGPAATYHKQVSLSHV"
BASE COUNT	356 a	803 c	736 g 439 t

Query Match 2.9%; Score 58; DB 77; Length 2334;
Best Local Similarity 59.9%; Pred. No. 5.31e-19;
Matches 175; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

Db 439 tgtgcgggcatcggctacaacctgacgacatgcccaaccagttccaacacgacacgc 498
|||||
401 TGTGCCACGGCATCGAATACCAGAACATCGGGCTGCCCAACCTGCTGGGCCACGAGACCA 460
QY
499 aggaacgaggcgggccctggagggtgcacacagtcttgcgcgctgggtggagatccaatgctcgc 558
|||||
461 TGAAGGAGGTGCTGGAGCAGCCGGCGCTTGGATCCCGCTGTCATGAAGCAGTGCCACC 520
QY
559 cggacctgcgcttcttcctatgcactatgtacacgcccattctgtctgtccccgactaccaca 618
|||||

QY.	521	CGGACACCAAGTTCCTGTGCTGCTGCTCTTCGCCCCCCCGTCTGCCCTCGATGACCTAGACG	580
D _b	619	agccgctgcgcgcctgcgcgctgggtgtgcgagcgcgcacaggccggctgctgcgcgcgtga	678
QY	581	AGACCATCCAGCCATGCCACTCGCTCTGCGTGCAGGTGAAGGACCGCTGCGCCCCGGTCA	640
D _b	679	tcgcgcagctacggcttcgcctcggcccgagcgcctgatgactgcgaccgcctccc	730
QY	641	TGTCCGCGCTTCGGCTTCGCCCTGGCCCGGACATGCTTGAGTGGACCGGTTTCC	692

RESULT	3	MMU43321	2421 bp	DNA	ROD	24-FEB-1996
LOCUS						
DEFINITION		Mus musculus	putative transmembrane receptor	(frizzled 8) gene,		

ACCESSION	U43321
NID	g1151259
KEYWORDS	.
SOURCE	house mouse.
ORGANISM	Mus musculus

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Murinae; Mus.

REFERENCE 1 (bases 1 to 2421)

AUTHORS	Wang, Y., Macke, J. P., Abella, B. S., Andreasson, K., Worley, P., Gilbert, D. J., Copeland, N. G., Jenkins, N. A. and Nathans, J.
TITLE	A large family of putative transmembrane receptors homologous to the product of the <i>Drosophila</i> tissue polarity gene <i>frizzled</i>
JOURNAL	J. Biol. Chem. 271 (8), 4468-4476 (1996)

JOURNAL
MEDLINE
J. Biol.
96224032

REFERENCE 2 (bases 1 to 2421)

AUTHORS

TITLE Direct Submission
JOURNAL Submitted (14-DEC-1995) Jeremy Nathans, Molecular Biology and Genetics, Johns Hopkins Medical School, 725 N. Wolfe Street, Baltimore, MD 21205

FEATURES	Location/Qualifiers
SOURCE	1-2421

```

SOURCE
1:12421
/organism="Mus musculus"
/chromosome="19"

```

```

/chromosome= 18
/map="between Tpl2 and Cdh2"
188 2245

```

```

CDS
188..2245
/gene="frizzled 8"
/note="putative transmembrane receptor"
/codon_start=1

```

```

/product="transmembrane receptor"
/db_xref="PID:g1151260"
/translation="MEWGYLEVTSLLAALAVLQRSSGAAAAASAKELACQEITVPLCK
GIGNYTYMPNQFNHDTQDEAGLEVHQFWPLVEIQCSDDLKFLCSMTPTICLEDYKK
PLPPCRSVCEKAKAGCAPLMROYGFAPDRMCRDLPEQGNPDTLCMDYNRTDLTTAA
PSPGRRLLPPPPGPEQQPSGSGHSPPGARPPIHRRGSSRGSDAAAPPSRGGKARPP
GGGAAPCEPGQCRAPMVSVSSERHPLYNRVKTQGIANCALPCHNPFSSQDERAFTVF
WIGLSVLCFVSTFATVSTFLIDMERFKYPERPIFLSACYLFVSGVSLVRLVAGHEK
VACSGGAPGAGRGGGAGGAAAAAGAGRAGGASSPGARGEYEELGAVEQHVRYETTGPAL
LCTVFFLLVYFFGMASIIWWVILSLTFWLAAGMKWNEIAGYSQYFHLAALWLPVSVK
SIAVLNLSVVDGDPVAGICVYGNQSLDNLRGFVLAPLVLYLFGTFELLAGVSLFRIT
RSVIKQGGGPTKTHKLEKLMIRLGLFTVLYTPAAVVACLFYEQHNRPWEATHNCP
CLRDLPQDARRPDYAVFMLKYFMCLVYVGITSGVWVWSGKLTLESWRLCTRCCWASKG
AAVGAGAGGSGPGGGGPGGGGGGGGGGSLYSDVSTGLTWRSGTASSVSYPKQMP
sov"

```

BASE COUNT	403 a	772 c	779 g	467 t
ORIGIN				

Query Match 2.8%; Score 56; DB 87; Length 2421;
Best Local Similarity 59.4%; Pred. No. 1.17e-17;
Matches 177; Conservative 0; Mismatches 121; Indels

Db 306 tgccgttggtgcaaggcatcggtttacaacattacacttaatgcccaaccagttcaaccacg 365
 |||||
 QY 395 TGCAGCTGTGCCACGGCATCGAATACCAGAACATGCGCTGCCCAACCTGCTGGGCCACG 454
 |||||
 Db 366 abacgcaagatgaggcgggtcttagaggtgcaccagttttggccgctggtggagatatacagt 425
 |||||

Ug	AGCAGT 514	16
	aggact 485	9
	ATGACC 574	11
	gcgcgc 545	
	GGGCCC 634	
	tgcc 603	7
	TCCC 692	

-JUL-1993
) homologue
ma cDNA to
ata;
pha; Muridae;
dley, M.S.,
ler, G.J. and

homologue"

3: Gaps 2:

acacgaa 287
|||
ACGAGAC 458
agtgtctc 347
|||||
AGTGCCA 518

ttg--ctgga 404
 || || ||
 ATGACCTAGA 578
 || || ||
 ggagggcgt 464
 || || ||
 GGCGCTCGGT 638
 || || ||
 tccc 518
 || || ||
 tccc 692

24-MAY-1996

CDS

E COUNT

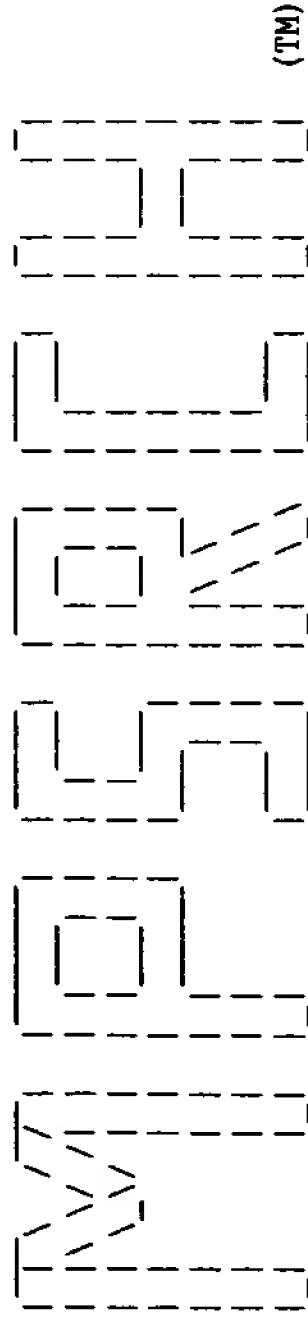
923; 3: G

gccaacagaa 326
||||| ||
ggccacgagac 458

tgcaagtgtc 386
||||| ||
agcagtgcca 518

--tg--ctgga 443
||| ||| ||
atgacctaga 578

gcgaagccct 503



(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPerch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 9 08:30:17 1997; MasPar time 741.29 Seconds
1311.186 Million cell updates/sec

Tabular output not generated.

Title: >848439-2-trans.ige
Description: (1-885) from translate.seq
Perfect Score: 734
N.A. Sequence: 1 ATGTYNCARGGCGGCGGWS.....SNATMGNARYTNCARTGY 885
Comp: TACRANGTYCCGCGGCGGWS.....WNTANKNTTYRANGTYACR

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 362067 seqs, 549138275 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: embl-new3
1:BCT 2:FUN 3:GEN1 4:GEN2 5:HTG1 6:HTG2 7:HUM 8:INV
9:ORG 10:MAM 11:VRT 12:PLN 13:PRO 14:ROD 15:SYN 16:UNC
17:VIR

Database: genbank99
18:BCT1 19:BCT2 20:BCT3 21:BCT4 22:BCT5 23:BCT6 24:BCT7
25:BCT8 26:BCT9 27:BCT10 28:BCT11 29:GEN1 30:GEN2
31:GEN3 32:HTG1 33:HTG2 34:HTG3 35:INV1 36:INV2 37:INV3
38:INV4 39:INV5 40:INV6 41:INV7 42:INV3 43:INV9 44:INV10
45:INV11 46:MAM1 47:MAM2 48:MAM3 49:VRT1 50:VRT2 51:VRT3
52:VRT4 53:PAT1 54:PAT2 55:PAT3 56:PAT4 57:PAT5 58:PHG
59:PLN1 60:PLN2 61:PLN3 62:PLN4 63:PLN5 64:PLN6 65:PLN7
66:PLN8 67:PLN9 68:PLN10 69:PLN11 70:PRI1 71:PRI2
72:PRI3 73:PRI4 74:PRI5 75:PRI6 76:PRI7 77:PRI8 78:PRI9
79:PRI10 80:PRI11 81:PRI12 82:PRI13 83:PRI14 84:PRI15
85:ROD1 86:ROD2 87:ROD3 88:ROD4 89:ROD5 90:ROD6 91:ROD7
92:ROD8 93:STR 94:SYN 95:UNA 96:VRL1 97:VRL2 98:VRL3
99:VRL4 100:VRL5 101:VRL6 102:VRL7 103:VRL8 104:VRL9
105:VRL10

Database: genbank-new3
106:BCT 107:GEN1 108:GEN2 109:HTG1 110:HTG2 111:INV
112:MAM 113:VRT 114:PHG 115:PLN 116:PRI1 117:PRI2
118:ROD 119:SYN 120:UNA 121:VRL

Database: u-emb150_99
122:part1

Statistics: Mean 15.331; Variance 11.948; scale 1.283

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description	Pred. No.
1	720	98.1	1776	85	D50462	Mouse SDF5 mRNA, comp	0.00e+00
2	89	12.1	2421	87	MMU43321	Mus musculus putative	1.25e-22
3	75	10.2	2334	77	HSU43318	Human putative trans	3.36e-16
4	64	8.7	1770	38	CEU43316	Caenorhabditis elegans	2.45e-11
5	63	8.6	2831	87	MMU43317	Mus musculus putative	6.61e-11
6	51	8.3	2085	40	DMU65589	Drosophila melanogaster	4.75e-10
7	59	8.0	2260	87	MMU43320	Mus musculus putative	3.36e-09
8	56	7.6	1913	38	CEU63557	Caenorhabditis elegans	6.05e-08
9	56	7.6	4540	90	RA1FRZ2H	Rattus norvegicus Dro	6.05e-08
10	54	7.4	1923	80	HUMFRIZ	Human frizzled gene p	4.04e-07
11	52	7.1	1912	90	RATFRZ2H	Rattus norvegicus Dro	2.64e-06
12	52	7.1	34576	8	CET23D8	Caenorhabditis elegans	2.64e-06
13	52	7.1	34576	38	CET23D8	Caenorhabditis elegans	2.64e-06
14	52	7.1	34576	111	CET23D8	Caenorhabditis elegans	2.64e-06
15	49	6.7	215	57	I28278	Sequence 5 from paten	4.18e-05
16	47	6.4	215	57	I28278	Sequence 5 from paten	2.54e-04
17	46	6.3	2624	87	MMU43205	Mus musculus frizzled	6.19e-04
18	44	6.0	354	8	OFU89259	Oxytricha fallax 57kD	3.59e-03
19	44	6.0	354	111	OFU89259	Oxytricha fallax 57kD	3.59e-03
20	43	5.9	354	111	OFU89259	Oxytricha fallax 57kD	8.51e-03
21	43	5.9	354	8	OFU89259	Oxytricha fallax 57kD	8.51e-03
22	39	5.3	74911	33	HS209H1	Human DNA sequence **	2.45e-01
23	38	5.2	370	8	OFU89262	Oxytricha fallax 57kD	5.52e-01
24	38	5.2	370	111	OFU89262	Oxytricha fallax 57kD	5.52e-01
25	37	5.0	1509	23	INSREP8	Insertion sequence IS	1.23e+00
26	37	5.0	1513	24	MHU01217	Mycoplasma hyorhinis	1.23e+00
27	37	5.0	1518	23	INSGDLA	Insertion sequence IS	1.23e+00
28	37	5.0	2187	43	PFASHP70H	Plasmodium falciparum	1.23e+00
29	37	5.0	30864	38	CET01H10	Caenorhabditis elegans	1.23e+00
30	36	4.9	2492	87	MMU43319	Mus musculus putative	2.70e+00
31	36	4.9	140571	34	HS46H23	Human DNA sequence **	2.70e+00
32	35	4.8	1291	11	XLU78598	Xenopus laevis Frzb p	5.86e+00
33	35	4.8	1291	113	XLU78598	Xenopus laevis Frzb p	5.86e+00
34	35	4.8	1375	52	XLU68059	Xenopus laevis frezzl	5.86e+00
35	35	4.8	4284	24	MHVLF4JC	M.hyorhinis vlpA, vlp	5.86e+00
36	35	4.8	33200	36	CEF54B3	Caenorhabditis elegans	5.86e+00
37	35	4.8	47745	33	CET26E3	*** SEQUENCING IN PRO	5.86e+00
38	34	4.6	294	83	HUMUT612B	Human STS UT612, 3' p	1.25e+01
39	34	4.6	413	44	PFU17812	Periplaneta fuliginos	1.25e+01
40	34	4.6	1054	39	DDU67923	Dictyostellium discoid	1.25e+01
41	34	4.6	3618	38	CEZYGI1	Caenorhabditis elegans	1.25e+01
42	34	4.6	4413	69	YSPCRM1N1	Yeast crml-N1 gene fo	1.25e+01
43	34	4.6	13815	88	MUSDYSA	Mouse dystrophin mRNA	1.25e+01
44	34	4.6	34731	78	HSU75931	Human Xp22 cosmid U25	1.25e+01
45	34	4.6	119704	63	OSCHLPLXX	O.sinensis chloroplas	1.25e+01

ALIGNMENTS

RESULT 1 D50462 1776 bp mRNA 19-DEC-1996
LOCUS Mouse SDF5 mRNA, complete cds.
DEFINITION

TITLE A large family of putative transmembrane receptors homologous to the product of the Drosophila tissue polarity gene frizzled

JOURNAL J. Biol. Chem. 271 (8), 4468-4476 (1996)

MEDLINE 96224032

REFERENCE 2 (bases 1 to 2421)

AUTHORS Abella,B., Wang,Y., Macke,J.P. and Nathans,J.

TITLE Direct Submission

JOURNAL Submitted (14-DEC-1995) Jeremy Nathans, Molecular Biology and Genetics, Johns Hopkins Medical School, 725 N. Wolfe Street, Baltimore, MD 21205

FEATURES

source Location/Qualifiers

1..2421

/organism="Mus musculus"

/chromosome="18"

/map="between Tpl2 and Cdh2"

188..2245

/gene="frizzled 8"

/note="putative transmembrane receptor"

/codon_start=1

/product="transmembrane receptor"

/db_xref="PID:g1151260"

/translation="MEWGYLEVTSLLAALAVLQRSAGAAASAKELACQEITVPLCK GIGYNYTMPNQFNHDTQDEAGLEVHQFWPLVEIQSPDLKFFLCSMYTPILEDYKK PLPPCRSVCEARAKAGCAPLMROYGFAMPDRMCRDLPEQGNPDTLICMDYNRTDLTTAA PPSPPRRLLPPPPGEGQPPSGSGHSRPPGARPPHRRGSGSGDAAAPPSRGGKARPP GGGAAPCEPGCCQCAPMVSSESRHPLYNRVKTGQIANCALPCHNPFPSQDERAFTVF WIGLMSVLCFVSTFATVSTFLIDMERFKYPERPIIFLSACYLFVSVGYLVRIVAGHEK VACSGGAPGAGCGGAGGAGAAAGACAGRASSPGARGEYEELGAVEQHVRYETTGPA LCTWVFLIVYFFGMASSIWWIILSTWFLAAGMKWGEAIIAGYSQVFHLAAILVPSVK STAVLALSSVDGDPVAGICVYVGNOSLDNLRGFVLAPLVIIYIFIGTMFLLAGFVSLFRI RSVIKQQGGPTTKHKLEKLMIRLGFTVLTYTPAAVVVACLFYEQHNRRPWEATHNCP CLRDLPDQARRPDYAVFMKLYFMCILVVGITSGVWVWSGKTLESWRALCTRCWASKG AAVGAGAGSGSGSGPGPGGGGGGGGSLYSVSTGLTWRSCTASSVSYPKQMPPL SQV"

BASE COUNT 403 a 772 c 779 g 467 t

ORIGIN

Query Match 12.1%; Score 89; DB 87; Length 2421;

Best Local Similarity 41.1%; Pred. No. 1.25e-22;

Matches 124; Conservative 47; Mismatches 131; Indels 0; Gaps 0;

Db 306 tgccgtgtgcaaggcatcggttacaaactacattacatgcccaaccagttcaaccacg 365
| | :| | :| | | | | | :| | :| | | | | | :| | :| | | | | | | | |

Qy 140 TNCARYTNTGYCAYGGNATNGARTAYCARAYATGMGNYTNCNAAYYTNYTNGNCAYG 199

Db 366 acacgaagataggcggcctagaggtgcaccagtttgccgctgglggagatacagt 425
| | | | | | :| | :| | :| | | | | | :| | :| | | | | | | | |

Qy 200 ARACNATGAARGARGCTNYTNGARCAGCNGCGNCTGGATNCCNYTNGTNAAGAART 259

Db 426 gctcccgacctaagtcttctgtgtagcatgtacacjcccatctgcctggaggact 485
| : :| | | :| | :| | | | :| | :| | :| | :| | :| | :| | :| |

Qy 260 GCAYCCNGAYACNAARAARTYYTNTGYWSNYTNTYGCNCCNGTNTGYTNGAYGAY 319

Db 486 acaagaagcctctgccgcttgcgtctgtgtgaacgcgcaagcgcgctgcgcgc 545
| | :| | | | | :| | :| | :| | :| | :| | :| | :| | :| | :| |

Qy 320 TNGAYGARACNATNCARCONTCYCAYWSNYTNTGYTNCARGTNAARGAYMGTGYGNC 379

Db 546 cgctcatgcgccagtagcgcttgccttgacctgacgcgatcgctgcgatcggttgcgg 605
| | | | :| | | :| | | | | | | | :| | | | | | :| | | | | |

Qy 380 CNGTNGTWSNGONTTYGNTTYCCNTGGCCNGAYATGTYTNGARTGYGAYMNTTYCCNC 439

Db 606 ag 607

Qy 440 AR 441

RESULT 3 HSU43318 2334 bp mRNA PRI 24-FEB-1996

LOCUS Human putative transmembrane receptor (frizzled 5) mRNA, complete cds.

DEFINITION

ACCESSION U43318

NID g1151251

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 2334)

AUTHORS Wang,Y., Macke,J.P., Abella,B.S., Andreasson,K., Worley,P., Cilbert,D.J., Copeland,N.G., Jenkins,N.A. and Nathans,J.

TITLE A large family of putative transmembrane receptors homologous to the product of the Drosophila tissue polarity gene frizzled

JOURNAL J. Biol. Chem. 271 (8), 4468-4476 (1996)

MEDLINE 96224032

REFERENCE 2 (bases 1 to 2334)

AUTHORS Abella,B., Wang,Y., Macke,J.P. and Nathans,J.

TITLE Direct Submission

JOURNAL Submitted (14-DEC-1995) Jeremy Nathans, Molecular Biology and Genetics, Johns Hopkins Medical School, 725 N. Wolfe Street, Baltimore, MD 21205

FEATURES

source Location/Qualifiers

1..2334

/organism="Homo sapiens"

/chromosome="2"

/map="2q33-34"

/tissue type="retina"

321..2078

/gene="frizzled 5"

/note="putative transmembrane receptor"

/codon_start=1

/product="transmembrane receptor"

/db_xref="PID:g1151252"

/translation="MARPPSAPPSLLLLLLAQLVGRAAAASKAPVCQEITVPMCRGI GYNLTHMPNQFNHDTQDEAGLEVHQFWPLVEIQSPDLRFFELCTMTPTICLPDYHKPL PPCRVCERAKAGCSPLMRQYGFAMPERSMCDRLPVLGRDAEVLCDMDYNRSEATTAPP RFPKAPTLPKPPGAPASGCEGACGPFVKCREFPVILKESHPLYNKVRTGQVPNC AVPCYQSFSADETFEATFWIGLMSVLCFISTSTTVATFLIDMTFRYPERPIIFLSA CYLCSVLGFLVRIVGHASVACSRHNHIHYETTGPAICTIVFLVYFFGMASSIWWI ILSLTFWFLAAMKMGNEAIIAGYQYFHAAWLIPIPSVKSIITALALSSVDGDPVAGICVY GNQNLSLRRFVLGPLYLVLTGLFLLAGFVSLFRIRSVIKQQGGTKDKLEKLMIRI GIFTLLYTPASIVVACLYEQHYRESWEAALTACPGHDTGQPRAKPEYWWIMLYF MCLVVGITSGVWVWSGKTVESWRRTSRCCRPRRGHKSGGAMAAGDYPEASAALTGR TGPPGPAATYHKQVSLSHV"

BASE COUNT 355 a 803 c 736 g 439 t

ORIGIN

Query Match 10.2%; Score 75; DB 77; Length 2334;

Best Local Similarity 39.7%; Pred. No. 3.36e-16;

Matches 116; Conservative 44; Mismatches 132; Indels 0; Gaps 0;

Db 439 tgtgccggcgcatcggtacaaactgacgcacatgcccaaccagttcaaccacgacgcg 498
| | | :| | | | | | :| | :| | | | | | :| | :| | | | | | | | |

Qy 146 TGTGYCAYGGNATNGARTAYCARAAYATGMGNYTNCNAAYYTNYTNGNCAYGARACNA 205

Db 499 aggacgagcgccctggaggtgcaccagttcttgccgctgggtgagagatccaatgctgc 558
| | | | :| | | :| | | | | | | | :| | | | | | :| | | | | |

Qy 206 TGAARGARGTNYTNGARCAGCNGCGNCTGGATNCCNYTNGTNAAGARTGYCAYC 265

ERP1IFLSGCTMTMSVAVIAGFVLQERVVCNERF SEDGYRTVVQGTKEGCTILFMML
YFFSMASSIWWVILSLTWFLAAGMKWGHEAIEANSQYFHAAWPAVKTITILAMQ
IDGDLISGVCVGLNSLDPLRGFVLAPLFVYLFIGTSLFAGFVSLFIRTIMKHDGT
KTEKLERLMVRIGVFSVLYTVPATIVIACYFYEQA FREHWSWSQHCKSLAIPCPA
HYTPRMSPDFTYVMIKYILMTLIVGITSGFWISGKTLHSWRKFYTRLTNSRHGETTV"
BASE COUNT 306 a 671 c 590 g 356 t
ORIGIN

Query Match 7.4%; Score 54; DB 80; Length 1923;
Best Local Similarity 39.7%; Pred. No. 4.04e-07;
Matches 119; Conservative 41; Mismatches 137; Indels 3; Gaps 2;

Db 263 tcccgctgtgcacggacatcgctacaaccagaccatcatgcccaaccttctggyccaca 322
| | :| | | : | | | | | : | | | | : | | | | : | | | | : | | | | :
Qy 140 TNCARYTNTGYCAYGGNATNGARTAYCAPAAYATGMGNYT ICCNAAYYTNTNGNCAYG 199
| | :| | | : | | | | | : | | | | : | | | | : | | | | : | | | | :
Db 323 cgaaccaggagcagcagcctagaggtgcaccagttctatccgctggtgaagtgccagt 382
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | : | | | | :
Qy 200 ARACNATGAARGARGTNYTNGARCARGCNGCNGTGGATNCCNTNGTNGATGAARCART 259
| | :| | | : | | | | | : | | | | | : | | | | | : | | | | : | | | | :
Db 383 gctcgcccggaactgcgcttcttctctgtgctccatgtacgcacccgctgtgcaccg-tg--c 439
| : | | | | | | | | | | | : | | | | | | | | | | | : | | | | : | | | | :
Qy 260 GYCAIYCCNGAYACNAAARAARTYYTNTGYWSNYTNTTYGCNCCNGTNGYTNAYGAYY 319
| | :| | | : | | | | | : | | | | | : | | | | | : | | | | : | | | | :
Db 440 tggacacaggccatcccgccgtgccgtctctatctgtgagcgcgcgccagggctgcgaag 499
| | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | : | | | | :
Qy 320 TNGAYGARACNATNARCCTNTGYCAYWSNYTNTGYGTNARGTNAARGAYMGNTGYGCNC 379
| | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | : | | | | :
Db 500 ccctcatgaacaagtctggtttccagtgcccgagcgccctgcgctgcgagcacttcccg 559
| | | | | : | | | | | : | | | | | : | | | | | : | | | | : | | | | :
Qy 380 CNGTNATGWSNGCNTTYGGNTTYCCNTGCCNCAYATGTYNGARTGYGAYMGNTTYCCNC 439

RESULT 11
LOCUS RATERZH 1912 bp mRNA ROD 16-JUL-1993
DEFINITION Rattus norvegicus Drosophila polarity gene (frizzled) homologue
mRNA, complete cds.
ACCESSION LO2530
NID g310114
KEYWORDS frizzled gene; homologue; polarity gene.
SOURCE Rattus norvegicus (strain Sprague-Dawley) osteosarcoma cDNA to
mRNA.

ORGANISM Rattus norvegicus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Rattus.
REFERENCE 1 (bases 1 to 1912)
AUTHORS Chan,S.D.H., Karpf,D.B., Fowlkes,M.E., Hooks,M., Bradley,M.S.,
Vuong,V., Bambino,T., Liu,M.Y.C., Arnaud,C.D., Strewler,G.J. and
Niasenson,R.A.

TITLE Two homologs of the Drosophila polarity gene frizzled (fz) are
widely expressed in mammalian tissues
J. Biol. Chem. 267, 25202-25207 (1992)
MEDLINE 93094228

FEATURES Location/Qualifiers
source 1..1912
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/cell_line="UMR106-01"
/cell_type="osteoblast"
/tissue_type="osteosarcoma"
79..1791
CDS /note="Drosophila polarity gene (frizzled) homologue"
/codon_start=1

/db_xref="PID:g310115"
/translation="MRARSALPRSLRLLLP LLLLPAAQPAQFHGKGISIPDHGFC
QPIISIPLCTDIAYNOTIMP NLLGHTNQEDAGLEVHQFYPLVKVQCSPELRFFLCMYA
PVCVTLEQALPPCRSICERARQGC EALMNKFGQWPERLRCEHPPRHGAEQICVGQNH
SEDGTPALLTAPP SGLQPGAGTGGPGGGGAPPRYATLEHDFHCPRLVKVPYSLSY
KFLGERDCAAPCEPARPDGSMFFSHHTRFARLWILTWSVLCCASTFFTVTTSLVAMQ
RFRYPERP1IFLSGCTMTMSVAVIAGFVLQERVVCNERF SEDGYRTVCGGTKEGCTI
LFMMLYFFSMASSIWWVILSLTWFLAAGMKWGHA AIEANSQYFHAAWPAVKTITI
LAMQIDGDLISGVCVGLNRLDPLRGFVLAPLFVYLFIGTSLFAGFVSLFIRTIM
KHDCTKTEPLERLMVALJVFSLVLYTPATIVIACYFYEQA FREHWSWSQHCKSLA
IPCPAHYTPTSPDFTVYMIKYILMTLIVGITSGFWISGKTLHSWRKFYTRLTNSRHG
ETTV"

BASE COUNT 299 a 662 c 566 g 385 t
ORIGIN
Query Match 7.1%; Score 52; DB 90; Length 1912;
Best Local Similarity 39.3%; Pred. No. 2.64e-06;
Matches 118; Conservative 41; Mismatches 138; Indels 3; Gaps 2;

Db 224 tcccgctgtgcacggacatcgctacaaccagaccatcatgcccaaccttctggyccaca 283
| | :| | | : | | | | | : | | | | : | | | | : | | | | : | | | | :
Qy 140 TNCARYTNTGYCAYGGNATNGARTAYCARAAYATGMGNYTNCNAAYYTNTNGNCAYG 199
| | :| | | : | | | | | : | | | | | : | | | | | : | | | | : | | | | :
Db 284 cgaaccaaggagcagcggcctggaggtgcataattctaccgcgctggtgaagtgccagt 343
: | | | | | : | | | | | : | | | | | : | | | | | : | | | | : | | | | :
Qy 200 ARACNATGAARGARGTNYTNGARCARGCNGCNGTGGATNCCNTNGTNGATGAARCART 259
| | :| | | : | | | | | : | | | | | : | | | | | : | | | | : | | | | :
Db 344 gctcgcccgagctgcgcttcttctctgtgctccatgtacgctcggtgtgcacgg-tg--c 400
| : | | | | | | | | | | | : | | | | | : | | | | | : | | | | : | | | | :
Qy 260 GYCAIYCCNGAYACNAAARAARTYYTNTGYWSNYTNTTYGCNCCNGTNGYTNAYGAYY 319
| | :| | | : | | | | | : | | | | | : | | | | | : | | | | : | | | | :
Db 401 tggagcagccatcccgccgctgcgctccatctgcgaacgcgcgcgccaaggtgcgagg 460
| | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | : | | | | :
Qy 320 TNGAYGARACNATNARCCTNTGYCAYWSNYTNTGYGTNARGTNAARGAYMGNTGYGCNC 379
| | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | : | | | | :
Db 461 cgctcatgaacaagttcggcttccagtgcccgagcgccctcgctgcgagcattcccc 520
| | | | | : | | | | | : | | | | | : | | | | | : | | | | : | | | | :
Qy 380 CNGTNATGWSNGCNTTYGGNTTYCCNTG-CCNGAYATGTYNGARTGYGAYMGNTTYCCNC 439

RESULT 12
ID CET23D8 standard; DNA; INV; 34576 BP.
AC Z81128;
NI el008304
DT 21-OCT-1996 (Rel. 49, Created)
DT 01-MAR-1997 (Rel. 51, Last updated, Version 5)
DE Caenorhabditis elegans cosmid T23D8
KW CD63 antigen like; Drosophila tissue polarity protein like;
KW histone H2A; histone H4; Mouse bright protein like;
KW transporter protein.
OS Caenorhabditis elegans
OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;
OC Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP 1-34576
RA Wild A.;
RT ;
RL Submitted (21-OCT-1996) to the EMBL/GenBank/DBJ databases.
RL Nematode Sequencing Project, Sanger Centre, Hinxton, Cambridge CB10
RL 1RQ, England and Department of Genetics, Washington University, St.
RL Louis, MO 63110, USA. E-mail: jes@sanger.ac.uk or
RL rw@nematode.wustl.edu
RN [2]

RX MEDLINE; 94150718.

RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J.,
RA Coulson A., Craxton M., Dear S., Du Z., Durbin R., Favello A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., McMurray A., Mortimore B.,
RA O'Callaghan M., Parsons J., Percy C., Rifken L., Roopra A.,
RA Saunders D., Shownkeen R., Smaldon N., Smith A., Sonnhammer E.,
RA Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaudin M.,
RA Vaughan K., Waterston R., Watson A., Weinstock L.,
RA Wilkinson-Sproat J., Wohldman P.;

RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of
RT C. elegans*;

RL Nature 368:32-38(1994).

CC Current sequence finishing criteria for the C. elegans genome
CC sequencing consortium are that all bases are either sequenced
CC unambiguously on both strands, or on a single strand with both
CC a dye primer and dye terminator reaction, from distinct
CC subclones. Exceptions are indicated by an explicit note.

CC

CC IMPORTANT: This sequence is NOT necessarily the entire insert
CC of clone T23D8. It may be shorter because we only sequence
CC overlapping sections once, or longer because we arrange for a
CC small overlap between neighbouring submissions.

CC The true left end of clone T23D8 is at 1 in this sequence.

CC The true right end of clone T23D8 is at 34576 in this sequence.

CC The true right end of clone T24D1 is at 12749 in this sequence.

CC Coding sequences below are predicted from computer analysis,
CC using the program Genefinder (P. Green, ms in preparation),
CC and other available information.

CC The start of this sequence (1..104) overlaps with the end of
CC sequence CET24D1.

FH Key Location/Qualifiers

FH

FT source 1..34576

FT /organism="Caenorhabditis elegans"

FT /clone="T23D8"

FT /chromosome="I"

FT complement(join(31876..31973,32210..32284,32331..32573,
FT 32640..33032,33192..33264))

FT /product="T23D8.8"

FT /note="protein predicted using Genefinder"

FT /note="Similarity to Mouse bright gene (TR:Q62431)"

FT /db_xref="PID:e304447"

FT /translation="MYELSDDVKKKEWLDWLNFMHRICKPVTRIPIMAKQVLDLYEL

Y

FT RLVVQHGLVEIINKLWREITKGLNPSSITSAFTLRTQYQKYLVDYECEKEKLSN

Q

FT SDLQQAIDGNRRREAPGRRTAPSPFLPFQLPHAASAAATMLNNQLNGLGMRNDLLDEN

T

FT LSLQASGLGCTSYGAEQMAILEAHQRNLERAQRAVQVCVARQSLGLTACSNNGNGNIH

N

FT SGRSTSSNDS DIPAKRPKLENDVKTNGASSMRISTKHSNPKTSMSVSMELNGITYQ

G

FT VLFALDETUSES"

FT complement(25597..25908)

FT /product="T23D8.5"

FT /note="protein predicted using Genefinder"

FT /note="Similarity to histone H4"

FT /db_xref="PID:e275927"

FT /translation="MSGRGKGGKGLCGGAKRRHKVLRDNIQGITKPAIRRLARRGGV

K

FT RISGLIYEETRGVLKVFLENVIRDVITYCEHAKRKTVTAMDVVYALKRQGRNTLYCGFG

FT CDS

FT complement(join(7258..7391,7494..7638,8798..8860,9211..934

3

FT ,10019..10101,10150..10212))

FT /product="T23D8.2"

FT /note="protein predicted using Genefinder"

FT /note="Similarity to Rat CD63 antigen (SW:P28648)"

FT /db_xref="PID:e276248"

FT /translation="MVEGGVTIVKYLLFLANLVLMVGGSLIIVGSLQLKFDNVLDI

L

FT GDERLATPILLLVIGSLCTLGLGCCGAIRENYCLTVSFAVLLALLITGMVRYHESR

G

FT VESAWDKTHQLFECCGVTNSSDWLFTTIPDSCCIEIEGCARENAPLFEFGCIHSVE

Q

FT WVLKNGAMVGGITCAVLAAIQLVGVCFAACLSKSLKDFHDFYY"

FT join(2786..3258,3758..4122,4187..4325,4866..5107,5220..534

0

FT ,5688..5906,6332..6479)

FT /product="T23D8.1"

FT /note="protein predicted using Genefinder"

FT /note="Similarity to Drosophila Tissue polarity protein

FT (TR:Q24760)"

FT /db_xref="PID:e175930"

FT /translation="MHRHILILFLGCLSADQLSSTSISSMNGFSTTRKCEHITIPM

C

FT KNLDYNQTVFPNLLGHTTQSEAGPAIAQFNPLIKVKCSEDIRLELCTVYAPVCTVLEK

P

FT IQPCRELCLSAKNGCESLMKPFQWPQDLCNKFVPTDLCVGKNSESSENSKSSND

V

FT TFGVSTIANEVVLSPKKCPHHMHTTSCSHFSLPLLSGRLPECSLTCEADNQVPMFEDG

R

FT VRRILRIWTAAMSVACFVCSLFTLVTLVDLSRFAYPVRPILYLAFCYLAISTVVMIG

V

FT VGEDGFACGTYGSTPTTLVTQGENVGCSALAVVHYFFFMSSCAWLVLCIAWFLAAN

L

FT KWGAESTIAALSPYFHAMCWGVPDAVLSVTVLVTVNSVDGVFTGICSVGNINPSALVYFF

F

FT TPIVVSLALGAVLLVCCGIWSMIRIRSIYIKIQHADVERNISKLEKLMIRICAFAIMYSL

P

FT TAMNAAIMWQAVNMPAWLEGWLHRCVRLQDRELFGFTYVPDDCPMDPKVAAPETIV

F

FT LLKYVSQLVVGITCAIWWVSSKTLSSYHKAYIALSSRPTVPAHVQDVNMNR"

FT 26277..26660

FT /product="T23D8.6"

FT /note="protein predicted using Genefinder"

FT /note="Similarity to histone H2A"

FT /note="cDNA EST yk72a2.3 comes from this gene"

FT /note="cDNA EST yk124d7.3 comes from this gene"

FT /note="cDNA EST yk124d7.5 comes from this gene"

FT /note="cDNA EST yk72a2.5 comes from this gene"

FT /db_xref="PID:e275250"

FT /translation="MSGRGKGGKAKTGKAKSRSSRAGLQFPVGRLHRILRKGNVAQR

V

FT GAGAPVYLAAVLEYLAAEVLELAGNAARDNKKTRIAPRHQLAVRNDDEELNKLACVT

I

FT AQGGVLPNIQAVLLPKKTGGDKE"

FT join(26949..27068,27120..27167,27232..27606,27661..28621,

FT 28689..28819,29049..29748,30273..30464,30510..30659)

FT /product="T23D8.7"

FT /note="protein predicted using Genefinder"

FT /note="cDNA EST yk116f3.3 comes from this gene"

FT /note="cDNA EST yk20f1.5 comes from this gene"

FT /note="cDNA EST yk116f3.5 comes from this gene"

FT /db_xref="PID:e275933"

FEATURES	Location/Qualifiers	
	source	
CDS	1..34576	
	/organism="Caenorhabditis elegans"	
	/clone="T23D8"	
	/chromosome="I"	
	join(2786..3258,3758..4122,4187..4325,4866..5107,5220..5340,5688..5906,6332..6479)	
	/note="protein predicted using Genefinder"	
	/codon_start=1	
	/product="T23D8.g"	
	/db_xref="PID:e275930"	
	/db_xref="PID:g1628244"	
CDS	/translation="MRRHILILFLGCLSDQRLSSTSISMMNGFSTTRKCEHITIPMCKNLQNTVPFNLLGHTTQSEAGPAIAQFNPLIKVKCEDIRLFLCTVTAPVCTVLEKPIQPCRELCLSAKNGCESLMKKEGFQWPDQDLCNKFPVTDLCVCKNSSESNSKKSSNDVTFGVSTIANEWLSPKCPHHMTTSGSHFSLPLLSGRLPECSLTCEADNQVPMWFDGRVRIRLRTWTAAWSVACFVCSLFTLVTFLVDSRFAYPVRPILYLAFCYLAISTVYMGVGEDGFACCTYGSTPTTLVTQGENVGCSALAVVHYFFFNSSCAWWLVLCIAWFLAANLKWGAESIAALSPYFHMCWGVPAVLSTVTLVTSVDGVTGICSVGNLPSALVYFFFTPTVVSIALGAVLLVCGIWSMIRIRSYIKLQHADVERNISKLEKIMLRICAFAIMYSIPTAMNAAIMYQAVNMPAWLEGWLHHRQVRLQDRELFGFTYPVDDCPMDPKVAAPEIIVFLIKYVSQVLVVGITCAIWWVSSKTLSSYHKAYLALSRSRPTVPAHVQVNMR"	
	complement(join(7258..7391,7494..7638,8798..8860,9211..9343,10019..10101,10150..10212))	
	/note="protein predicted using Genefinder"	
	/codon_start=1	
	/product="T23D8.f"	
	/db_xref="PID:e276248"	
	/db_xref="PID:g1628243"	
	/translation="MVEGVTVIKYLLFLANLVLWVGGLSLIVGSLIQLKFDNVLDILGDERLATPILLVIGSLCTLGFLGCCAIRENYCLTVSFAVLALLITGMVRYHESRGVESAWDKTHOLFECCTGVTNSSDMLTFTTIPDSCIEIEGCARENAPLFEPGCTHSVEQWVLKNGAMVGGICAVLAAIQLVGCFACCLSKLSILKDFHDFY"	
	complement(join(10723..10767,10978..11142))	
	/note="protein predicted using Genefinder"	
CDS	/codon_start=1	
	/product="T23D8.e"	
	/db_xref="PID:e275929"	
	/db_xref="PID:g1628242"	
	/translation="MGVTKPFEPKKNKISNFSPLPFFVLLVRKNKKNTTAMHQACRFFYFKKIKMLIAKGRIQGIKARPI"	
	complement(join(11188..11658,11707..11823,12477..12917,13249..13431))	
	/note="protein predicted using Genefinder"	
	/codon_start=1	
	/product="T23D8.d"	
	/db_xref="PID:e275928"	
CDS	/db_xref="PID:g1628241"	
	/translation="MGRKAFIDKKASQSFRLVDPDNRTSRERFKPTQEHLEEQQYGVYYDDYDYLOHMRAVNEPMKLENVHEVEKTTIKTSGSSAFPPAPPLFGLVGALKKPEFFDEDVANALEEVTDDNTGELEDNFITLAGGLIDERTTVYSTRSTRGEDSEEDDDDEDDEMYDDYNDDELFGEEAVGEIRVERADQVIDNAFEELMDREYNTDQIGELDGDYDVGGALEPNAGRLHLKLINDKGPNAEYDEEAKHYVVRMRILIEEGVIKOEYEIVEVDEGTNKKMKWDGESFATQYTNIYHPTLIKEPRLSKALKRFDKAVEEMDIAEDEDDEDEMEDDDDKESVFSTVSTFRPKNETPEQRSIRKKAVKEARKLRRVEKKANKTMAFEKRRKIVFIHSY"	
	13855..14139	
	/note="protein predicted using Genefinder"	
	/codon_start=1	
	/product="T23D8.h"	
	/db_xref="PID:e275931"	
	/db_xref="PID:g1628245"	
	/translation="MGETIEKWIFRRILVAFEDDDVADHLEIPNFLAFLVLVLHVLL	
CDS	complement(25597..25908)	
	/note="protein predicted using Genefinder"	
	/codon_start=1	
	/product="T23D8.b"	
	/db_xref="PID:e275927"	
	/db_xref="PID:g1628239"	
	complement(25597..25908)	
	/note="protein predicted using Genefinder"	
	/codon_start=1	
	/product="T23D8.c"	

CDS	GIHQILSAYLLAARSFLETTIILDNHHVQAAPEFVHCFRQEFVCQTS"	
	complement(join(13883..14240,14292..15409,15456..16341,16423..16656,16705..16807,17668..17721,17774..18278,18330..18656,18888..19113,19170..19705,19751..19828,19884..19991,20041..20353,21385..21569,21857..21931))	
	/note="protein predicted using Genefinder"	
	/codon_start=1	
	/product="T23D8.c"	
	/db_xref="PID:e276247"	
	/db_xref="PID:g1628240"	
	/translation="MNSTGEQRGAGPYHHPMPYHVDIPTTSHPSWHQAPHGGLPTPPYHPQQHPNYISMQPPPSQQQOMTPOALSTQQQQVQQQQQRLYSSPSRGPAAQAQNKFERTEQWNNQAMPYSQSPAPPVAPSVMSYHQDDRMMLSVNTTWTNQFPDSQRCYSSNGSTCENVNPEMMQMNITQAAEQAIRLWENTPDPMQRLHMAKTIRTWIRQKF AQVDQANMPCVQQILNI IIDGLKPQVQLPISYYAQLWYNLLDILRRFTFLPIISPYIHQVVQMFCPRENGPQDFRELICNLISLNWQKDPHMKHCANQVFQIFNCIIMGVKNKELRTEFAQHLKFEKLVGTLSYFNQPVHPMINPAIFIIFRFIISKDTRLKDYFIWNNNPHDQPPPTGLIIKLANAVMIGSYRLIAGQNPELTPQNELAHLLQVIRTFDLGLLLHSDAIDGCFVRSDBGCAITTVVQYPNNDLIRAGCKLLLVQVSDAKALAKTLENILPFLLRLLIEIHPDDEVISGTCFLSNVVAHQVKDIAIRSNAILLHTIISKYPRLDLTDAPKRRVCEIICNCLATJNNFLMMWIPTPNGETKTAGNEKQQVCKFIEIDILKKLMSCLSCEMDTPGLLELRSTILRSFILLRTPFVPKDGVLNVIDENRKENLGHICAAYSWFRQPNNTRTQSTKQQLVERTISLLVIMEQCGAEVAQSYSIDCPNLNLNGQVKPTFIHNVLVVCDKILEHCPTRADIWTIDRPMLEGLTNHRNSDIKAANSILLSRPENDLLAAMSRFFHAKEDSDTSSEDEVEDQKVNKSAKFRDDLDFMAGPEDEKRVVVRARQDKKFEDELGIIKQNRDAKSNKDLNRLLTGFDSLAKAYDKSKTVFQRQNVANPREYIRSLVEIEDYVYVKLWDDKDAKSALSNNAKALPPLRQKLKYYIKDQQLQDLVTDYRNPDEDGYETPEDEDDDDDFGEVSESKAESPGRPSEKAAVSDSDSDSDSSSNWSSEPNSSDDEDSVTKMEQLRRYFLKKEFRVESKDDKKEKKRVIRVKEAVEEDDDADWTPVNERKSVVHFDKKEEVDHVMIKKLNEMVSARGKRTTDRNQHVANLRKLLLEVSEKELGGLGINVKISFCIISALFELNAKISDHMEYETFTTILQTVNSLLDLLIGTDRVKLSVTYAEEDENLKDDTQEVRIQGSILIAVQRLDGLAKILQNADCHSNDIETIKAKKMDMCSLIEKAEYVELRNDSGIFDKHEVCVKYMMRIEHAYYKQDQNEEDACKLMDYLCNKIYTLDDKRLRQRAMLCHVYYLAVHDKWHRARDLLMSHMQAIVDHSVDVTQILNRTICQLGLCAFRHGFIREAHQGLSEIQNTQRAKELLAAQVGTQRHEKTAEQEKIDRSRQVPYHMHINVELMECVYLICSMILLEIPHMASCEFEMRRRMLSRSPHYQLKQSEKASLTGPPENTREHVVAASKAMLNGDMKKCQDYIVNDKMNQKVNLFHNAETVKGMVVRRIQEESLRYYL LTYSTVYATVSLKKLADLFELSKDVHSIISKMIIQEELSATLDEPTDCLIMHRVEPSRLQMLALNLSDKLQTLAENNEQILEPTRGRGGYQGPSPGRNERQGDQKQKSGSGYQGERRGQGQDCKRGNWGSQGGQRRHPQKPRAF"	
	join(14867..15084,15127..15265)	
	/note="protein predicted using Genefinder"	
CDS	/codon_start=1	
	/product="T23D8.i"	
	/db_xref="PID:e276249"	
	/db_xref="PID:g1628246"	
	/translation="MSFANETVSESAETELADSSWVEDLSVDIRMINDSLHVTHKKKISGAVPLVMNSQIVNMAEHSTLSQSFLIINILLILILIFVVGMLDAAHHVHPANFVFIEVARIVTKFNVLILGLFNI"	
	15774..16091	
	/note="protein predicted using Genefinder"	
	/codon_start=1	
	/product="T23D8.j"	
	/db_xref="PID:e275932"	
CDS	/db_xref="PID:g1628247"	
	/translation="MSNFFLGIKVNNRLITINWSPIRIIVFLDSSFDSNLSLLFFFLVIFALHAEFFFOEIPTKLLHLSHTVFVIRRIIGFWFRPVRGVIVAVTVTIAHCSFLRWFSR"	
	complement(25597..25908)	
	/note="protein predicted using Genefinder"	
	/codon_start=1	
	/product="T23D8.b"	
	/db_xref="PID:e275927"	
	/db_xref="PID:g1628239"	
	complement(25597..25908)	
	/note="protein predicted using Genefinder"	

CDS

26277..26660

/note="protein predicted using GeneFinder"

/codon_start=1

/product="T23D8.k"

/db_xref="PID:e276250"

/db_xref="PID:g1628248"

/translation="MSGRKGKGKAKTGKAKSRSSRAGLQFPVGRHLHRIIRKGNVAQR

VGAGAPVYLAIVLEYIAAEVLELAGNAARONKKTRIAPRHQLAVRNDEELNKLLAGV

TIAQGGVLPNTQAVLLPKKTGDK"

join(26949..27068,27120..27187,27232..27606,27661..28621,

28689..28819,29049..29748,30273..30464,30510..30659)

/note="protein predicted using GeneFinder"

/codon_start=1

/product="T23D8.l"

/db_xref="PID:e275933"

/db_xref="PID:g1628249"

/translation="MEDQWLLSAIYDDDLVEKIKVRSSSTSSRSTINVPSELENEFLSS

SSGSRVSDLLYLPHEENREPKLIGKPLPSTTGRFLSLIANHFQITCGSIHGYI

RFDPDIPSKLNRITLRLQEQNPGLIECPLVFDGIHTVYSTELINKEVNNSVINVA

...
Note: remainder of annotations omitted.

Query Match 7.1%; Score 52; DB 38; Length 34576;
Best Local Similarity 40.1%; Pred. No. 2.64e-06;
Matches 120; Conservative 44; Mismatches 130; Indels 5; Gaps 5;

Db 2910 ttccaatgtgcaaaaatctggattacaatcaaacagatttccaatcttctcgacata 2969

| | : | ||: | | || ||: | | | : | || ||||: | | || ||: | | || ||:

QY 140 TNCARYTNTGCAYCGGNATNGARTAYCARAAYATGMGNTNCCNAAYTYTNTGNCAYG 199

Db 2970 caacacaatctgaagctgggtccagcaattgcgcaattccaatccaattaaagtta-aa 3028

:|| ||||: | | ||||: | | ||||: | | ||: | | ||: | | ||: | |

QY 200 ARACNATGAARGARG-TNYTNGARCARGCGNGCNGTCGATNCCNYTNGTNATGAARCAR 258

Db 3029 tgctcagaagatatctgtctcttcttctgtactgtctatgcacctgtctgtacag-t-ac 3086

||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:

QY 259 TCYAYCCNGAYACNARAARITYTYTNTGYWSNTYNTTYGCNCCNGTNTGYTNGAYGAY 318

Db 3087 t-cgaaaaaccaattcaaccatgtcgaagaattgtgttatctctgcaaaaaatggatcgag 3145

: || | : || || |||| ||||: || ||||: | | ||||: | | ||||: | | ||||:

QY 319 YTNGAYGARACNATNCARCCNTGCYAYWSNYTNTGYGTNCARGTNAARGAYMGNTGYCN 378

Db 3146 tcattaatgaaaaagtttggaattcaatggccagatcaattggattgtaacaaattccc 3204

| | ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:

QY 379 CCNGTNATGWSNGCNTTYGGNTTYCCNTGCCCNCAAYATGYTNGARTCGAYMGNTTYCC 437

RESULT 14

LOCUS CET23D8 34576 bp DNA INV 01-MAR-1997

DEFINITION Caenorhabditis elegans cosmid T23D8.

ACCESSION 281128

NID g1628237

KEYWORDS CD63 antigen like; Drosophila tissue polarity protein like; histone H2A; histone H4; Mouse bright protein like; transporter protein. Caenorhabditis elegans.

SOURCE Caenorhabditis elegans

ORGANISM Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 34576)

AUTHORS Wild,A.

TITLE Direct Submission

JOURNAL Submitted (21-OCT-1996) Nematode Sequencing Project, Sanger Centre, Hinxton, Cambridge CB10 1RQ, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail: jes@sanger.ac.uk or rw@nematode.wustl.edu

REFERENCE 2 (bases 1 to 34576)

AUTHORS Wilson,R., Ainscough,R., Anderson,K., Baynes,C., Berks,M., Bonfield,J., Burton,J., Connell,M., Copsey,T., Cooper,J., Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A., Fulton,L., Gardner,A., Green,P., Hawkins,T., Hillier,L., Jier,M., Johnston,L., Jones,M., Kershaw,J., Kirsten,J., Laister,N., Latreille,P., Lightning,J., Lloyd,C., McMurray,A., Mortimore,B., O'Callaghan,M., Parsons,J., Percy,C., Rifken,L., Roopra,A., Saunders,D., Showkeen,R., Smaldon,N., Smith,A., Sonhammer,E., Staden,R., Sulston,J., Thierry-Mieg,J., Thomas,K., Vaudin,M., Vaughan,K., Waterston,R., Watson,A., Weinstock,L., Wilkinson-Sproat,J. and Wohldman,P.

TITLE 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans

JOURNAL Nature 368 (6466), 32-38 (1994)

MEDLINE 94150718

COMMENT Current sequence finishing criteria for the C. elegans genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note. IMPORTANT: This sequence is NOT necessarily the entire insert of clone T23D8. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions. The true left end of clone T23D8 is at 1 in this sequence. The true right end of clone T23D8 is at 34576 in this sequence. The true right end of clone T24D1 is at 12749 in this sequence. Coding sequences below are predicted from computer analysis, using the program GeneFinder (P. Green, ms in preparation), and other available information. The start of this sequence (1..104) overlaps with the end of sequence CET24D1.

FEATURES

source location/Qualifiers

1..34576

/organism="Caenorhabditis elegans"

/clone="T23D8"

/chromosome="I"

join(2786..3258,3758..4122,4187..4325,4866..5107,5220..5340,5688..5906,6332..6479)

CDS

/note="protein predicted using GeneFinder; Similarity to Drosophila Tissue polarity protein (TR:Q24760)"

/codon_start=1

/product="T23D8.l"

/db_xref="PID:e275930"

/db_xref="PID:g1628244"

/translation="MHRHILILFLGCLSDQRILSSSTISSMNGFSTTRKCEHITIPMCKNLDTYVTFPNLGHHTTQSEAGPAIAQFNPLIKVKGSEDIRLFLCTVYAPVTVLEKPIQPCRELCLSAKNGCESLMKFGQWPDQLDCNKFVTDLCVKGNSSESSNKKSSNDVTFGVSTIANEVVLSPKKCPHHMHTTSGSHSLPLLSGRLPECSLTCEADNQVPMFDRVRRILRIWTAAWSVACFVCSLFTLVTLVDTLSRFAYPVRPTILYAFCYLAISTVYMGVVGEDGFACGTGSTPTTLVTQGENVGCALAVVHYFFFMSSCAWMLVLCIAMFLAANKWGAESIAALSPYFHAMCWGPVAVLSVTVLVTNSVDGVFTGICSVGNLNPVALVYFFFTPIVVSALGAVLVCGIWSMIRISYIKLQHADVERNISKEKLMRLIGAFAIMYSLPTANAAIMYQAVNMPAWLEGLHHRVCVRLQDRELFGTYPVDDCPMDPKVAAPETIIFLLKYVSQLWGITCAIWWVSSKTLSSYHKAYLALSSRSPTVPAHVQWNMR"

complement(join(7258..7391,7494..7638,8798..8860,9211..9343,10019..10101,10150..10212))

/note="protein predicted using GeneFinder; Similarity to Rat CD63 antigen (SW:P28648)"
/codon_start=1
/product="T23D8.2"
/db_xref="PID:e276248"
/db_xref="PID:gl628243"
/translation="MVEGGVTVKYLFLANLVLVWGGSLIIVGSIQLKFDNVLDI LGDERLATPILLVIGSLCTLLGLGCCGAIRENYCLTVSFAVLLALLITGMVRYHES RGVESAWDKTHOLFECGVTNSSDMLTFTTPDSCCIEIEGCARENAPLFEPCGCIHS VEQWVLKNGAMVGGICAVLAAIQLVGVCFACCLSKSIILKDFHDFYY"
complement(join(10723..10767,11212..11658,11707..11823, 12477..12917,13249..13431))

CDS

/note="protein predicted using GeneFinder; cDNA EST cm16f9 comes from this gene; cDNA EST CEESF93R comes from this gene; cDNA EST CEESF93FC comes from this gene"
/codon_start=1
/product="T23D8.3"
/db_xref="PID:e304448"
/db_xref="PID:gl865750"
/translation="MGRKKAFIGDKASQSFRLVPDNRSESRERKPTQEHLEEQKYGV YYDDYDYLQHMRAVNEPMKLENVHEEVEKTTIKTSGSAFPAPPLFGLVGALKKPE FFDEDVANALEEVTDDRNTGELEDNFITLAGLLDERTTVYRSTRGEDSEEEEDDDE DDEMYYDNDDELFGEEAVGEIRVERAQORVIDNAFEELMDREYNTDQIGELDCDDYD VGGALEPNAGRLHLKLNKGPNSAEYDEELAKHYVRERMRLLIEEGVIKDEEYEIVEV DEGTNKKMKWDCESFATQYTNINYHPTLIKEPGLSRKALKRFDKAVEEMDIAEDED DDEDMEDDDDKESVFTSVSTFRPKNETPEQRSLRKKA VKEARKLRRVEKKANKTWF AE EKRLAKGRIGQIKARPI"
complement(join(13883..14240,14292..15409,15456..16341, 16423..16656,16705..16807,17668..17721,17774..18278, 18330..18656,18888..19113,19170..19705,19751..19828, 19884..19991,20041..20214))

CDS

/note="Similarity to Yeast nuclear transport protein (SW:P32497); cDNA EST yk44d4.5 comes from this gene; cDNA EST yk33b1.5 comes from this gene; cDNA EST yk8e5.5 comes from this gene; cDNA EST yk8e7.5 comes from this gene; cDNA EST yk4b7.5 comes from this gene; cDNA EST yk30g3.5 comes from this gene; cDNA EST yk8e5.3 comes from this gene; cDNA EST yk8e7.3 comes from this gene; cDNA EST yk44d4.3 comes from this gene; cDNA EST yk37b1.3 comes from this gene; cDNA EST yk37d3.3 comes from this gene"

/codon_start=1
/product="T23D8.4"
/db_xref="PID:e304461"
/db_xref="PID:gl865751"
/translation="MSYHQGDRMSMLSNTMTNQFPDSQRCYSNGSTCENVNPEM MQMNTQAAEQAIRLWENTPDPQRLHMAKTIRTWLQDKFAQVDQANMENCVQIILN I IYDGLKPQPVQLPISYYAQLWYNLLDILRRFTFLPIISPYIHQVQMFCPRENGPQD FRELIGNLISLNWQKDPHKHCANQVQIFNCIIMGVKNKELRTEFAQHILKFEKLVTG LSEYFNQVHPGMINPAIFIIFRFIISKDTRLKDYFIWNNNPHDQPPPTGLIILKNA VMIGSVRLIACGNPETLPQNPELAHLIQV IIRTFD LLLGLLHDSDAIDCFVRS DVGVA ITTVQYPNNDLIRAGCKLLQLQVSDAKALAKTPLENLPFLRLRIEHPDEVIYSGT GFLSNVVAHQHVQDAIRNSAIFLLHTIISKYPRLDELTPADKRNRCVCIICNCLRT LNNFLMMWIPTPNGETKTAGPNEKQOVCKFIEIDILKKMLSCSLCEGMDTPGLLEIRS TILRSFILLRTPFPVKDGLVNIDENKENLGHICAAYSWFVFPQPNRTRTQSTKQQ LVERTISLLLMEQCCAEKEVAQYSYSDCPNLNLNGNQVKPTFIHNVLVWCDKILE HCPTRAD IWTIDRPMLEGLTNHRNSD IAKAANSLLSRFPENDLLAAMSRFHAKEDSD DAKSNKD LNRLLTGFD SLAKAYDKSTVFQRQNVANPRFVIRSLVEIEDIYVNLKWDK DAKSALSKNNAKALPPLRQKLKKYIKDQQLDVLVTDYRVNPDDEGYETPEDEDDDFG EVSESKAESPCKPSEKAAVSDSDSDSDSSNWSSESPESNDSDEDSVTKMEQLRR YFLKKEFRVESKDDKKEKKRVIRVKEAVEEDDADWTPVNRKSVVHFDPKEETHD VMIKKLINEVMSARGKRTTRDNRQHVANLRKLLLEVSEEEKELGLGINVKISFCIISALFEL

NAKISDHMEYETMTTIQTVNSLLDLIGTRVKLSVTYAEEDENLKDDTQEYRIQGS ILIAVQRLDGELAKILQNADCHSNDYIEKLIKAEKDMCSLIEKAEKYVELRNDSGIFDK HEVCKVYMMRIEHA YKYQDQNEEDACKLMDYLCNKIYTLDDEKRLQRAMLCHVYYL AVHDKWHRARDLLLSHMQAIVDHSVDVTQILYNRTICQLGLCAFRHGFIREAHOGILS EIQNTQRAKELLAQAVGTRQHEKTAEQEKIDRSRQVPYHMHINVELMECVYLICSMLL EIPHMASCEFEMRRRMLSRSPHYQLKQSEKASLTGPPENTREHVVAASKAMLINGDWKK CQDYIWDKNQKQVWNLFHNAETVKGMVRRIQEESLRTYLLTYSTVYATVSLKKIAD LFLSKDKVHSIISKMIQEELSATLDEPTDCLIMHRVEPSRLQMIALANLSDKLQTLA ENNEQILEPTRGRGGYQCPGSWFPGRNERQGDQKQSGCYQGERRGGQGDGKRGNWG SQGGQRRHFQKPRAF"

CDS

complement(25597..25908)
/note="protein predicted using GeneFinder; Similarity to histone H4"
/codon_start=1
/product="T23D8.5"
/db_xref="PID:e275927"
/db_xref="PID:gl628239"
/translation="MSGRKGKGLGKGAKRHRKVLRDNIQGITKPAIRRLARRGGV KRISGLIYEETRGVIKVLENVIRFAVTYCEHAKRKTVTAMDVVYALKRQGRTLVGF G"

CDS

26277..26660
/note="protein predicted using GeneFinder; Similarity to histone H2A; cDNA EST yk72a2.3 comes from this gene; cDNA EST yk124d7.3 comes from this gene; cDNA EST yk124d7.5 comes from this gene; cDNA EST yk72a2.5 comes from this gene"
/codon_start=1
/product="T23D8.6"
/db_xref="PID:e276250"
/db_xref="PID:gl628248"
/translation="MSGRKGKAKTGKAKSRSSRAGLQFPVGRLHRILRKGNVAQR VGAGAPVYLAAVLEYLAAEVLELAGNAARDNKKTRIAPRHLQLAVRNDEELNKL LAGV TIAQGGVLPNIQAVLLPLPKKTGGDK E"
join(26949..27068,27120..27187,27232..27606,27661..28621, 28689..28819,29049..29748,30273..30464,30510..30659)
/note="protein predicted using GeneFinder; cDNA EST yk116f3.3 comes from this gene; cDNA EST yk20f1.5 comes from this gene; cDNA EST yk116f3.5 comes from this gene"

CDS

/codon_start=1
/product="T23D8.7"
/db_xref="PID:e275933"
/db_xref="PID:gl628249"
/translation="MEDQWLLSAIYDDDLVEKLVRSSTSSRSTSINVPSIENEFLSS SSGSRVSDDLYLHP IENREPFKLGKPLPSTTGRFLSLLANHFQITCNGSIIHQYVI RFDPDIPSKKLNRTILRTLQEQNPGLIECPLVFDGIHTVYSTELINVKEVNSVINVA GVNTKESPNIKFkLYLTHVDSFLDWTkIITGNQDQNKRLMMHAIDTVFRQTSTGNFH AVLQSFSSIAQNSA ILP JHGLGWGTN LGVREV CYQNVVETFDTLTMNLDVATT TFYRPVALVEFLAEIILEVPLATVTDGRSLSDVQKKFNREVAGLKVETRHCSCPRRFR VARCTWKPTENISFH SETAGNQDSKPLSLVEYKRRYNIDLTYKHLPCIEVGRTEC ILPLELCYVVSQRCIKKLINEQQIANLIRATSRNATERQNVAWSLQNR LKMDNDVNAV KFG LKVEAQLKIEGRVLPVPRLLYRSPN LKQECVTPVNNGTWDMRGKNFYSGIQR EWAIVCFASPEIIGEASMRSFVRNLNVASEIGMPFLEEHRFCRYAEPDQTVKLLLEHL NEQYNLQLVLCIVPGKSVVYGELKRKGLLGLTQCVRSONVSKASPH TSLNLCMKIN SKLGINVILSSPPQSLNSEPVLFIGCHLTRSLASSDSTSSIAHCDS SIACLVGSM DGHP TQFSP IFTQPRHQRTIVDMCEMTREAIINF RKSTGFKPHKII IYRAGIADTVV DEIMQTELRAVRDACAMIEGFPQGITF IGLDVTHHTRFLFAANEKDRVGN SQNV PAGT LVETGLTVNNLFEYVLVSHAGIQGTSRPTKYVVMWDDNS IPSADIHEMTYQLCHTQSR CTRSVSIPSPVYAKLVAQRAKILMADENFDMERFLCVSWF"
complement(join(31876..31973,32210..32284,32331..32573,

...

Note: remainder of annotations omitted.

Query Match 7.1%; Score 52; DB 111; Length 34576;
Best Local Similarity 40.1%; Pred. No. 2.64e-06;

Dec 9 09:05

848439-2-trans,trans

Matches 120; Conservative 44; Mismatches 130; Indels 5; Gaps 5;

Db 2910 ttccaatgtgcaaaaatctggattacaatcaaacagtatctccaaatcttctcgacata 2969

QV 140 TNCARYTNTGYCAYGGNATNGARTAYCARAAAYATGMGNYTNCNCAAYYTNYTNGNCAYG 199

Db 2970 caacacaatctgaagctggtccagcaattgggcaattcaatccattaatattaaagtta-aa 3028

QV 200 ARACNATGAARGARG-TNYTNGARCAFGCNGGNCNTGGATNCCNYTNGTNGTATGAARCAR 258

Db 3029 tgctcagaagataatcgctctctcttcttctgtactgtctatgcacctgtctgtacag-t-ac 30866

QV 259 TGYCAYCCNGAYACNAARAARTTYTNTGYWSNYTNTTYGCNCCNGTNTGYTNGAYGAY 318

Db 3087 t-cgaaaaaccaattcaaccatgtcgagaattgtgttatctgcataaaatggatcgag 3145

QV 319 YTN GAYGARACNATNCARCCNTGYCAYWSNYTNTGYGTNCARGTNAARGAYMGNTGYGCN 378

Db 3146 tcattaatgaaaagtgttgattccaatggccagatcaattggattgtaacaaattccc 3204

QV 379 CCNGTNA TGWSNGCNTTYGGNTTYCCNTGCCNGAYATGYTNGARTGYGAYMGNTTYCC 437

RESULT 15

LOCUS	I28278	215 bp	DNA	PAT	30-OCT-1996
-------	--------	--------	-----	-----	-------------

DEFINITION Sequence 5 from patent US 5569830.

ACCESSION I28278

NID
a181901

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 215)

AUTHORS Bennett, A., Labavitch, J.M., Powell, A. and Stotz, H.

TITLE Plant inhibitors of fungal polygalacturonases and their use to control fungal disease

JOURNAL Patent: US 5569830-A 5 29-OCT-1996;

FEATURES

source 1..215

/organism="unknown"

BASE COUNT	15 a	8 c	25 g	26 t	141 others
ORIGIN					

Query Match 6.7%: Score 49; DB 57; Length 215;

Best Local Similarity 12.1%: Pred. No. 4.18e-05:

Matches	25;	Conservative	81;	Mismatches	98;	Indels	2;	Gaps	2;
---------	-----	--------------	-----	------------	-----	--------	----	------	----

Db 11 vvertascndakakkdgnntsswttdccnrtgwcddtttyrvnnndsgnhkysanyvvgg 70
:: :: |::: : :| : :|::: : :|::: : :|::: : :

293 ARNWSRCANARRAAYTTYTNGTRTCNGGRTGRCAYTGYTTCATNACNARNNGNATCCAN 234

Db 71 nnvgaakthyyhtnvgadskvtvdsynasgtssnggtdgnrsgadsgyskta mter 130

Cp 233 GCNCCNG-CYTGTCNARNACYTCYTTCA TNGTYTCRTGNCNARNARRTTNGGNARNCK 175

Db 131 nrtgktannavdsrnmgdasvsgskntkhhaknsadgk-vgsknmngdnrnyrgtgtksnv 189

Cp 174 CATRTTYGRTAYTCNATNCCTGRCANARYTGNARRTTNGCNGNATNGGYTTRCARTT 115

Db 190 anncaankrdvgavannkccasct 215

[illegible]

Co 114 NWSNCKYTTTRTANWSRAARTCNGGYT 89

Dec 9 09:05

848439-2-trans.18e



Search completed: Tue Dec 9 09:05:38 1997

Job time : 2121 secs.

Dec 9 10:15

848439-3-trans.rng

1

WQSEF (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Tue Dec 9 10:11:03 1997; MasPar time 92.45 Seconds
931.301 Million cell updates/sec

Tabular output not generated.

Title: >848439-3-trans
Description: (1-825) from translate.seq
Perfect Score: 688
N.A. Sequence: 1 WSGCGMGGNGNYTYYT.....SNATMGNARYTNCARTGY 825
Comp: SWCGKNCNCRANARRA.....WNTANKNTTYRANGTYACR

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 142080 seqs, 52183452 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq28

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29

Statistics: Mean 11.254; Variance 12.824; scale 0.878

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	122	17.7	1047	2 Q10572	Human Natriuretic Pep	3.70e-29
2	99	14.4	1047	2 Q10572	Human Natriuretic Pep	3.46e-21
3	53	7.7	501	3 N50030	Sequence encoding new	5.08e-06
4	53	7.7	501	3 N50027	Sequence encoding new	5.08e-06
5	53	7.7	565	6 Q35072	HCV envelope region n	5.08e-06
6	52	7.6	501	3 N50033	Sequence encoding new	1.04e-05

Dec 9 10:15

848439-3-trans.rng

2

c	7	51	7.4	498	3 N50034	Sequence encoding new	2.11e-05
c	8	51	7.4	501	3 N50023	Sequence encoding new	2.11e-05
c	9	51	7.4	501	3 N50028	Sequence encoding new	2.11e-05
c	10	50	7.3	501	3 N50031	Sequence encoding new	4.27e-05
c	11	50	7.3	501	3 N50023	Sequence encoding new	4.27e-05
c	12	50	7.3	501	3 N50029	Sequence encoding new	4.27e-05
c	13	49	7.1	501	3 N50024	Sequence encoding new	8.62e-05
c	14	49	7.1	501	3 N50032	Sequence encoding new	8.62e-05
c	15	49	7.1	501	3 N50025	Sequence encoding new	8.62e-05
c	16	49	7.1	501	3 N50024	Sequence encoding new	8.62e-05
c	17	49	7.1	984	17 Q94336	Degenerate Alteromona	8.62e-05
c	18	48	7.0	501	3 N50026	Sequence encoding new	1.74e-04
c	19	47	6.8	498	3 N50034	Sequence encoding new	3.48e-04
c	20	47	6.8	501	3 N50031	Sequence encoding new	3.48e-04
c	21	47	6.8	501	3 N50029	Sequence encoding new	3.48e-04
c	22	47	6.8	501	3 N50032	Sequence encoding new	3.48e-04
c	23	47	6.8	501	3 N50028	Sequence encoding new	3.48e-04
c	24	45	6.5	501	3 N50027	Sequence encoding new	1.38e-03
c	25	45	6.5	501	3 N50030	Sequence encoding new	1.38e-03
c	26	45	6.5	501	3 N50033	Sequence encoding new	1.38e-03
c	27	45	6.5	565	6 Q35072	HCV envelope region n	1.38e-03
c	28	43	6.2	501	3 N50026	Sequence encoding new	5.42e-03
c	29	42	6.1	204	1 N81164	Base substituted E.co	1.06e-02
c	30	41	6.0	204	1 N81164	Base substituted E.co	2.08e-02
c	31	41	6.0	657	7 Q43519	Degenerate FMN reduct	2.08e-02
c	32	40	5.8	91	9 Q51746	Oligonucleotide probe	4.05e-02
c	33	39	5.7	91	9 Q51746	Oligonucleotide probe	7.85e-02
c	34	39	5.7	501	3 N50025	Sequence encoding new	7.85e-02
c	35	39	5.7	774	3 Q22981	Gelonin toxin DNA	7.85e-02
c	36	39	5.7	984	17 Q94336	Degenerate Alteromona	7.85e-02
c	37	39	5.7	1065	17 Q94335	Degenerate Alteromona	7.85e-02
c	38	39	5.7	3871	2 N71302	HSV-1 gB and surround	7.85e-02
c	39	39	5.7	3871	2 N71302	HSV-1 gB and surround	7.85e-02
c	40	38	5.5	657	7 Q43519	Degenerate FMN reduct	1.51e-01
c	41	37	5.4	711	15 Q88282	V.fischeri flavin red	2.89e-01
c	42	37	5.4	1065	17 Q94335	Degenerate Alteromona	2.89e-01
c	43	35	5.1	420	14 Q88233	Bovine aFGF synthetic	1.04e+00
c	44	34	4.9	711	15 Q88282	V.fischeri flavin red	1.96e+00
c	45	33	4.8	997	3 Q22702	Sequence encoding Pla	3.66e+00

ALIGNMENTS

RESULT	1
ID	Q10572 standard; DNA; 1047 BP.
AC	Q10572;
DT	09-APR-1991 (first entry)
DE	Human Natriuretic Peptide Receptor B.
KW	NPBB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW	hyperaldosteronism; glaucoma; guanyl cyclase.
OS	Homo sapiens.
FH	Key Location/Qualifiers
FT	Peptide 1..22
FT	/label= signal sequence
FT	Protein 12
FT	/label= mature NPBB
FT	Domain 23..455
FT	/label= extracellular domain
FT	/note= "binds natriuretic peptides A, B and C]"
FT	Domain 456..456
FT	/label= transmembrane domain
FT	Domain 479..1047
FT	/label= cytoplasmic domain
FT	/note= "GC and protien kinase activity"

[illegible]

RESULT	8	
ID	N50023	standard; DNA; 501 BP.
AC	N50023;	
DT	04-SEP-1991	(first entry)
DE	Sequence encoding new modified human beta interferon polypeptides	
DE	IFNX 416.	
KW	Antiviral; cell growth regulator; immune system regulator;	
KW	antiproliferative; ss.	
OS	Homo sapiens.	
FH	Key	Location/Qualifiers
FT	CDS	1..501
	/*tag= a	
PN	EP-163993-A.	
PD	11-DEC-1985.	
PF	17-MAY-1985; 105750.	
PR	17-MAY-1984; GB-012564.	
PA	(SEAR.) SEARLE G D & CO.	
PI	Bell LD, Boseley PG, Porter AG;	
DR	WPI; 85-311944/50.	
DR	P-PSDB; P50022.	
PPT	New modified human beta interferon polypeptide(s) - prepd. by	
PPT	plasmid transformed bacteria, with improved antiviral,	
PPT	anti-proliferative and immune regulating actions	
PS	Claim 28; Chart 2a, page 32; 71pp; English.	

CC	Compared with interferon beta prepd. by recombinant methods, the
CC	INFs of the invention are more active and have different affinities
CC	for cell surface receptors (allowing selective targetting); they
CC	have higher therapeutic index; improved stability against microbial
CC	breakdown during synthesis; and better in vivo solubility and
CC	stability. They are also easier to recover from incubation mixts.
SQ	Sequence 501 BP; 107 A; 31 C; 69 G; 80 T;
	Query Match 7.4%; Score 51; DB 3; Length 501;
	Best Local Similarity 16.0%; Pred. No. 2.11e-05;
	Matches 50; Conservative 95; Mismatches 167; Indels 1; Gaps 1;
Db	166 gcnymbacnathtaygaratgtytbcaraayathttgcnathttymgdcargaywnnnhn 225 : : :: :: : :: : :: : :: :
QY	410 SNWSNGAYCAYTYNTCCNGCNACGARGCGNCNAAAGCTNTGYGARGCNTGYAARA 469 : : :: :: : :: : :: : :: :
Db	226 wnnnacngntggaaaygaracnathgtngaraayybytbgcnaaygn-taycaycarat 284 : . : : : :: :: ::: :: : : ::: :
QY	470 AYARAAAYCAYGAYAAAYATNATGGARACNTNTGYAARAAAYCATTYGCNYTNA 529 : : :: :: : :: : :: : :: :
Db	285 haaycayytbaatacngtnybtbgargataarytbgaraargargayttyacnmddgnaa 344 :: : : :: : : :: : :: : :: : :: :
QY	530 ARATNAARGTNAARGARATNACNTAYATNAAYMGNCAYACNARATNATNTNGARACNA 589 : : :: :: : :: : :: : :: :
Db	345 rytbatgwinnwnnytbcayytbaarmgdtaytagcnmgdathytbcaytayytbaargc 404 : : : : : :: : : :: : :: : :: :
QY	590 ARWSNAARACNATNTAYAARYTNAAAYGGNGTNSNGARMGNCAYYYTNAARAARWSNGTNY 649 : : : : : :: : : :: : :: : :: :
Db	405 naargattaywnncaytgycntggacnathgtnmgdgtngarathybtmgaaytyta 464 : : : : : : :: : :: : :: :
QY	650 TNTGGYTNAARGAYWSNYTNCARTGYACNTGYGARGARATGAAYGAYATNAAYGCCNCCNT 709 : : : : : : :: : :: : :: :
Db	465 yttyathaaymgd 477 ::: : : :
QY	710 AYYTNGTNTATGGG 722 ::: : : :

```

RESULT 9
ID N50028 standard; DNA; 501 BP.
AC N50028;
DT 04-SEP-1991 (first entry)
DE Sequence encoding new modified human beta interferon polypeptides
DE IFNX 445.
KW Antiviral; cell growth regulator; immune system regulator;
KW antiproliferative; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1..501
FT /*tag= a
PN EP-163993-A.
PD 11-DEC-1985.
PF 17-MAY:1985; 105750.
PR 17-MAY-1984; GB-012564.
PA (SEAR ) SEARLE G D & CO.
PI Bell LD, Boseley PG, Porter AG;
DR WPI; 85-311944/50.
DR P-PSDB; P50027.
PT New modified human beta interferon polypeptide(s) - prepd. by
PT plasmid transformed bacteria, with improved antiviral,
PT anti-proliferative and immune regulating actions
PS Claim 28; Chart 2f, page 37; 71pp; English.
CC Compared with interferon beta prepd. by recombinant methods, the
CC INFs of the invention are more active and have different affinities
CC for cell surface receptors (allowing selective targeting); they

```


Cp 248 CANACNGGCGCRAANARNWSRCANARRAAYT 218